In traditional schools there were nevertheless, fairly strict institutional limits to substitution of inputs. There is a trend in modern pedagogy to break up those rigidities and the question then becomes how and to what extent this is possible. A complicating factor in education, as opposed to production for the market, is that we have to deal with substitutability in the production function of one individual, and in the aggregate production function of say, a school, where substitution may occur across individual pupils.

Of special interest for this study is the elasticity of substitution between students' time and other inputs. Would it be possible, for example, to decrease student time by using more expensive materials, like programmed books, or by using more teacher time, and would it pay off by lowering costs? Investigations into educational production functions in order to determine elasticities of substitution between inputs would seem to be a very desirable research project.

If we instead look at the related problem of marginal returns to an isolated change in one of the inputs, we are equally ignorant, although common sense would lead us to believe that those returns are diminishing. It would, for example, seem to be plausible enough that ultimately there is diminishing marginal returns to the student time input for students of any given ability (or other initial traits). It seems to be safe to assume that other kinds of inputs, like teacher time, space, equipment, and materials also show diminishing returns with respect to individual students as well as in the aggregate. In fact some of them, such, for example, as school space

will very rapidly reach a limit where the marginal product in all probability is zero if not negative.[1]

Summary

In this chapter we reviewed briefly some attempts to arrive at estimates of educational production functions. We discussed measurement difficulties and in particular those caused by use of proxies for inputs and problems of collinearities among variables. A case was made against the assumption of technological efficiency and it was concluded that policy actions cannot be based on estimates of educational production functions until we know this assumption to be at least approximately valid.

We related production maximization to cost minimization and found that by the underlying assumptions regarding production the latter approach, in fact, also takes technological efficiency for granted. However, we were able to push the analysis somewhat forward by tying efficiency to time to completion, asserting that procedures are efficient to the extent that time to "mastery" is minimized given the learner and the state of technological knowledge. The introduction of resource limitations makes the problem an economic one.

---

[1]For a thorough discussion of marginal returns to inputs into education see Thomas, The Productive School, pp. 63-78.

CHAPTER VI

SEQUENCE ANALYSIS

If in a sequence the learning of one step affects the learning of subsequent steps, and in particular if this changes time to completion in later steps, it is evident that costs of steps thus related must also be interrelated. In this chapter we will incorporate into the theory the dynamic elements introduced by a multi-step analysis. This is all the more necessary since learning tasks in schools are very seldom isolated from each other. In fact, the interrelations are often quite intricate, including both branching and simple chains, sometimes crossing over subject boundaries not to mention even more complicated patterns.

However, for simplicity we will discuss only simple chains of steps, one and only one step at each level being dependent on prior equally unique steps; in other words, no parallel steps exist within the sequence. First we will introduce an additive cost model and then a programming model will be presented.

The Cost of a Sequence of Steps

When for any given input the time to completion of a specified learning task is minimized for a particular student, we have obtained efficiency in the technical sense, something which is the concern of research in educational methodology. We will assume

this obvious condition for an economic optimum to be satisfied; in other words, all cost measures in the following analysis are supposed to be minimum under given input conditions. There are a number of questions that could be asked, and hopefully answered, by an economic analysis. One such question is to determine the most economical input combination. Another, which we have already discussed in connection with the basic cost model, is the determination of how much output should be produced.

When the problem is to bring a given number of students through a defined step sequence a new implemental variable is introduced, since by influencing learning at one step we also have some control over learning as an input into later steps.

Assuming for simplicity that all students achieve on the same level when entering the sequence, we may define output as mastery to a specified criterion on a test covering the whole sequence. We are now interested in finding the least cost "treatment conditions" that would lead us to the specified goal. There are many ways of varying the inputs in order to achieve this such as organizing the content of the steps and the order of sequencing, choosing among different methods of teaching, and trying out different kinds of equipment and materials. As long as students reach the specified final performance level, even mastery levels for prior steps may be thought of as variable. Variation in those mastery levels may be of some importance; Block has shown that varying mastery level requirements may influence time-to-completion patterns cumulatively over a sequence of steps. Maintaining a high level of mastery, he

found, could increase students' efficiency of learning, so that time required in subsequent steps for feedback and correction procedures would tend to go down as compared to time used for those purposes by students who were required to attain a lower mastery level.[1]

The sequencing of mastery level criteria should, therefore, be treated as a policy variable, since it may have a significant influence on time-to-completion patterns, and probably also on the cost of materials that have to be used. Even though clear empirical evidence is lacking, one certainly cannot exclude the possibility that imposing a very high mastery level for certain key steps in a sequence actually could lower time to completion over the whole sequence for some or all students.

We assume now that a given student body has to be brought through a given sequence of steps (n students and m steps). Let the cost of bringing the i^{th} student through m consecutive steps in a sequence be

$$(1)\quad C_i = S_{i1} + A_{i1}t_{i1} + \cdots + S_{im} + A_{im}t_{im}$$

$$= \sum_{j=1}^{m} [S_{ij} + A_{ij}\, t_{ij}]$$

S_{ij} = Cost of supplies (books, etc.) for student i in step j.

A_{ij} = Per time-unit cost of student i for step j (not counted in S_{ij}). This includes costs of teacher and student time, of school space, and so on.

---

[1]Block, "The Effects of Various Levels of Performance on Selected Cognitive, Affective, and Time Variables," pp. 69-75.

t_{ij} = Number of time units used by student i to complete step j. Summing over n student yields:

$$(2) \quad \sum_{i=1}^{n} C_i = \sum_{i=1}^{n} \sum_{j=1}^{m} [S_{ij} + A_{ij}\, t_{ij}]$$

Since sequences may take considerable time to complete it is no longer satisfactory to ignore discounting, and we therefore introduce a positive interest rate r by which costs can be discounted to present values. For simplicity r is assumed to be constant for all time periods. The present value of C_i is taken as of the start of the sequence.

We assume, as usual, complete knowledge of time distributions, but knowing the ordering of the students is no longer necessary. This knowledge may be thought of as arrived at from experience or by educational experiments. In practical applications of the theory the parameters will simply be ex post measures from teaching methods to be evaluated. Present value of C_i as per the start of the sequence would be

$$(3) \quad \hat{C}_i = \sum_{j=1}^{m} (S_{ij} + A_{ij}\, t_{ij})\ (1 + r)^{-q_j}$$

where q_j is <u>elapsed</u> time from the start of the sequence up til step j is finished. This time measure should not be confused with t, which refers to number of time periods actually spent at work in school. It should also be noted, however, that q normally depends on t; that is, $q = f(t)$. If, for example a subject is taught in school 5 periods each week and a particular student completes his task in t = 50 periods,

then q = 10 weeks or $\frac{10}{52}$ year. Therefore, if some ways are found by which time to completion for some or all students could be shortened the returns would depend both on the decrease in t and the associated decrease in q; there is a direct effect on outlays for time inputs and an indirect one on their calculated present values.

We now sum over students to get

$$(4)\quad \sum_{i=1}^{n} \hat{C}_i = \sum_{i=1}^{n} \sum_{j=1}^{m} (S_{ij} + A_{ij}\, t_{ij})\ (1 + r)^{-q_{ij}}$$

where q_{ij} obviously is elapsed calendar time from the start through step j for student i.

A change in the input mix or in mastery requirements may result in parameter changes ΔS_{ij}, ΔA_{ij}, Δt_{ij} and Δq_{ij}. After such a change (or such changes) (4) may be written

$$(5)\quad \sum_{i=1}^{n} \hat{C}_i + \Delta\hat{C}_i = \sum_{i=1}^{n} \sum_{j=1}^{m} [S_{ij} + \Delta S_{ij} + (A_{ij} + \Delta A_{ij})\ (t_{ij} + \Delta t_{ij})]\ (1 + r)^{-(q_{ij} + \Delta q_{ij})}$$

For costs to be minimized over the sequence, meaning that $\sum_{i=1}^{n} \hat{C}_i$ is a minimum, would require that

$$(6)\quad \sum_{i=1}^{n} \Delta C_i \geqslant 0 \quad \text{for all possible changes in input mix or}$$

mastery level requirements given that the goal is to have all or a specified fraction of the students pass the whole sequence and reach the final mastery performance level.

To exemplify the procedure: Take an increase in mastery requirements over one or more steps. Assuming no change in S (cost of books, etc.) or in A (costs per unit of time), this change would pay off if changes in t_{ij} are such that

$$\sum_{i=1}^{n} \sum_{j=1}^{m} (S_{ij} + A_{ij}\, t_{ij})\ (1 + r)^{-q_{ij}} >$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} [S_{ij} + A_{ij}\, (t_{ij} + \Delta t_{ij})]\ (1 + r)^{-(q_{ij} + \Delta q_{ij})}$$

There may, of course, be a cost increase ΔC and/or ΔA, associated with a change in mastery level requirements, in which case a positive pay off would be obtained if

$$\sum_{i=1}^{n} \sum_{j=1}^{m} (S_{ij} + A_{ij}\, t_{ij})\ (1 + r)^{-q_{ij}} >$$

$$\sum_{i=1}^{n} \sum_{j=1}^{m} [S_{ij} + \Delta S_{ij} + (A_{ij} + \Delta A_{ij})\ (t_{ij} + \Delta t_{ij})]$$

$$(1 + r)^{-(q_{ij} + \Delta q_{ij})}$$

In either case (6) is not satisfied; there exists a change (mastery level increase) that would make $\sum_{i=1}^{n} \Delta \hat{C}_i < 0$, and the costs summed up over the sequence could be lowered.

It would seem that the effect of decreasing costs by increasing mastery standards, when at all possible, could most easily be secured by imposing an amount of overperformance early in the sequence, because then the pay off period will be longer. There may be some negative side effects to watch out for unfortunately. Block found in his study that although maximal cognitive learning was produced (in his sample) by demanding a 95 per cent mastery level, students required to maintain an 85 per cent level scored somewhat higher on tests of interest and attitude.[1] The possibility of requiring a higher level of mastery on some key steps (most likely at the beginning of the sequence rather than towards the end) is an interesting alternative that has not yet been investigated.

In Chapter I we discussed overlearning as a source of the development of basic abilities.[2] What Block observed was not an increase in learning originating from a gain in some of those abilities; that possibility could be safely excluded, since his data originated from a three step sequence taught altogether during one week in school. Nevertheless the analysis of an extended sequence, where those abilities are indeed developed, would be perfectly parallel.

When evaluating educational projects the above relatively simple cost accounting should be quite easy to apply. It is important,

---

[1] Ibid., pp. 80-86.

[2] Basic abilities, it will be recalled, were defined as general traits of an individual, that are fairly enduring and relating to a variety of human tasks, in contrast to skills that pertains to performance on a specific task or a limited group of related tasks (Chapter I, pp. 10-11).

though, to note carefully the logic involved in condition (6). Knowing $\sum_{i=1}^{n} \Delta \hat{C}_i$ to be negative (costs have been lowered) is sufficient for concluding that an improvement has been achieved only if the benefits from taking the sequence have not thereby been decreased.

The mastery learning assumptions, however, make it possible for us to determine direction, if not the size, of an eventual benefit change. For any student whose time to completion over the sequence has been diminished benefits have actually been increased provided that the gain in time could be utilized. The reason for this is that he has made an opportunity gain; his accumulation of learning is the same, which is the signification of the mastery learning assumptions, at the same time as by finishing earlier he will have the opportunity to increase his benefits either by being gainfully employed at an earlier date or by being able to spend time on some other school activity, such as another subject or entering a new sequence in the same subject. If we are indifferent as to which students decrease their time the same holds also in the aggregate—in other words if time summed over students has gone down. Finally, knowing that over the full sequence $\sum_{i=1}^{n} \Delta \hat{C}_i \geq 0$ is not enough to conclude that we must be worse off to incur the extra C, since we could view this sequence as an early step in a still larger sequence—and so on and on. Thus far, indeed, we have not specified the full conditions for optimization.

Sketch of a Programming Model

To develop a programming model of timing in education is a study in its own right and quite an extensive one at that. Nevertheless we will in this section make an attempt to outline the major features of such a model, since it will provide us with some additional information of considerable value. The development of a detailed programming model of timing in education, in my opinion a desirable piece of educational research, is not intended.

We have already pointed out the different reasons for wanting to interfere with the distribution of costs over the students. The MC model assumes time distributions to be given and determines costs as a function of these distributions. A more general model would allow us to have time implicitly determined.

There are, as we recall, two major possible determinants of distributing costs: 1) Higher investment in some students may be preferable for economic reasons, because a larger pay off might be expected from this investment. These students are not necessarily those who show the most rapid learning. 2) Society may want to undertake compensatory education even if not economically profitable. So far we have not described how optimal conditions could be found when these two requirements are entered. Only the first one will be thoroughly analyzed, since the necessary extensions for the second are rather obvious.

In order to proceed in that direction we must make explicit several of the implicit functional relationships in the MC model.

The S parameter is a summation of costs of books, supplies, and similar cost figures. A and $g(x)$[1] include costs of time inputs, such as teacher and student time, space, equipment, administration. In a programming model these quantities will be explicitly included in a simultaneous equation system. The value of students' time in the MC model has to be predetermined by using foregone earnings. This is an approximation of a hypothetical market value of student time. It is less satisfactory for several reasons; if a mean value is used, which seems inescapeable, it may grossly overstate some while underestimating other students' time. Maybe even more important: it could be argued that the value of student time should reflect what the students could do in school, not what they could possibly do outside school. In the programming model student time will be formulated as a constraint and thus evaluated at its shadow price.

One of the problems that the MC model was said to be capable of contributing towards solving, was the question of how many students should be admitted into an education at various steps. This may however turn out to be of less importance in a situation where this decision is largely left to the individuals. (The society may, however, still want to have some guidelines as to whether to encourage students to take, or discourage them from taking a certain education.) Often it seems to be more realistic to treat the number of students and their entering characteristics, relevant for learning, as given

---

[1] In equation (8'), p. 49.

factors and "optimize" learning over this student population. For this a programming model would seem to be more appropriate.

The MC model assumes resources to be completely acquirable in necessary and predetermined proportions. This may be less realistic in education; the programming model will distinguish between purchased inputs and inputs that exist in given amounts; in addition the amount of money that could be spent is restricted by a budget constraint. The measure of MC is an ex post one; solely by assuming a special kind of self-selection where students present themselves in the proper order or by using instruments predictive of learning rates, could one expect students to be introduced into the system approximately by order of their time to completion of future learning tasks. At present such prediction is not very accurate; moreover serious questions could be raised against using such measures for selective admittance, even if the predictive power were very high.

The important time-cost relationship, is sufficient reason, nevertheless, for developing such prediction instruments and of investigating further the economic significance of timing in education. In the following programming model time to completion is one among many other inputs in the "production of learning," and thus independent of the ordering of the students with respect to time to completion of steps.

For the programming model we assume, as we did for the cost accounting present value model, that students have to be taken through a sequence of ordered steps with no steps left out. The problem for the educational "production plant" is to bring as many students as

possible as far as possible in some specified sense. On a scale appropriate to the macro point of view benefits may depend on the size of the output, but it is assumed that the range of possible outputs envisaged by the micro unit will be too small perceptibly to influence unit benefit values. Our earlier assumption that benefits are given may, therefore, be interpreted to mean that weights are given to each step in the sequence; in other words, there exist "relative prices" telling the school what "as many as possible as far as possible" would mean. The problem would then be to allocate existing resources in such a way as to maximize "total revenue."

For example, the weights may be signaling to the schools that efforts should be directed towards bringing as many students as possible through the lower steps, or they may imply that schools should invest heavily in a few students, bringing them very far even at some sacrifice for the slow learners.

It may, however, not be very realistic to assume that such an evaluation of steps has been effected, and especially not if we are dealing with rather small steps, as will usually be the case. Benefits may be attached to larger units than we are considering, and we will have to assume that steps within the sequence are homogeneous with respect to benefits, meaning that we are indifferent as to where on the scale an increase takes place. We arrive at our benefit measure by interpolation, as it were. In order to justify the procedure, however, we must make the steps equivalent in some sense, so that they can be added and compared. What we need, in other words, is a cardinal scale; by assigning benefit weights (prices) to the

steps we had such a scale. If we do not have prices, it will be critical how we measure the step units in terms of their learning content.

Suppose that a benefit value is given to a sequence of n steps. We then try to approximate the unknown benefit measure for each step by subdividing the value for the whole sequence in such a way that to each step will be attached a value that is proportional to its learning content, somehow measured. Unfortunately there seems to be no proper yardstick for this measurement task. We simply have to assume that there exists such a thing as amount of learning, and that this quantity can be subdivided in learning tasks of equal or comparable length.

Carroll points out that in the case of programmed instruction learning rate as measured by number of frames per time period, covered by a particular student, is remarkably constant from lesson to lesson, and that the same holds for increments in new vocabulary and grammar points in foreign language learning.[1] In some cases, then, physical units seem, at least approximately, to be sufficient. Carroll suggests also that the amount of learning measured in physical units need not always be a linear function of time. The learning curve for new words to be spelled increases at a decreasing rate, and "insightful" learning curves may show sudden, sharp increases after periods of very slow improvement.

The problem may, in fact, be more easy to solve in practice than in theory in that we have the possiblity of making steps equal by

---

[1]Carroll, "Problems of Measurement," in Mastery Learning, ed. by Block, pp. 37-41.

"construction," applying whatever measure of equality seems best to fit the situation at hand. Steps are equal because we decide that they are. It may even be possible to turn around Carroll's discussion of the possible shapes of the function relating learning and time and ask, what scale would make learning a linear function of time. We construct in other words, a scale, such that the probability is maximized that a given learner under the same learning conditions will use the same amount of time for any two successive steps.

In the programming model to be developed, we assume that we have a "learning content" scale so that steps are either equal or measurable in equivalent units, hence comparable and additive. We, furthermore, assume that we are indifferent as to where on the scale an increase takes place. The introduction of benefit weights (prices) applicable directly to the step measure would not change anything in principle with respect to optimal conditions or conclusions drawn, but it would make the mathematical notation considerably more complex.

The objective function in the programming model to be developed is assumed to be continuous and differentiable. Otherwise no restrictions on its shape are assumed except the seemingly very reasonable ones that it shows diminishing or at the most constant returns with respect to isolated changes in any one of the inputs and constant returns to scale. The function, therefore, is assumed to be concave over the set of values in which we are interested.

The characteristics of the learners are given. The schools, in other words, have no way of selecting their student body and thereby increasing output. Measure of output will be the added number of

student-steps taken in a given sequence.[1] No steps can be skipped but students may enter anywhere in the sequence, except that it is assumed that there are always new steps to be taken, no matter how fast or how advanced a particular student may be.

Outputs and inputs are flows, that is, amounts per production period, for example, teacher hours or student hours per week. This is, of course, the usual approach, but it seems worth pointing out nevertheless since some of the inputs are time inputs, and therefore to be interpreted as number of time units (e.g., hours) per production period (e.g., semester) hence time units per time period, which may be somewhat confusing if not made explicit.

In the following, subscripts will be used to refer to individuals, whereas superscripts will refer to inputs (resources). For example: t_i^j means the amount of time resource j used up by student i.

In the hope that the exposition will be easier to follow we start with the hypothetical case in which we have only one student in the system, or students of one type only, always taking the same amount of time to complete a step. Even from this oversimplified version of the model some implications may be drawn, but it is, of course, the generalization to n students that will yield the most interesting conclusions.

Inputs are:

$t^j = (t^1, \cdots, t^1)$ time inputs of which one, say t^1, is student time. $c^k = (c^1, \cdots, c^m)$ purchased inputs like books, and materials.

---

[1] If steps are equal in learning content; otherwise they have to be weighted by learning content.

Among time inputs are space, equipment, teacher time, and other school services made available to the students in certain amounts per time unit. They all exist in fixed amounts and cannot be increased in the short run. Depending upon existing conditions such as the time span of the analysis and conditions in the labor markets (especially for teachers) one or more of these inputs may be considered in the c-vector, as being purchaseable. One input can never be purchased, however, namely student time; the t-vector, therefore, always includes at least the element t^1. All c^k, $k = 1, \cdots, m$, are subject to a budget contraint; a limited amount of money is made available per production period.

Suppose now that as a starting point for the analysis, we make the assumption that all school services, whether purchaseable or not, are made available to the students in the same amounts per production period for each student. Suppose further that for a given student the production function to be maximized is:

$$(7)\quad \text{Max } 0 = f(t^1, \cdots, t^1, c^1, \cdots, c^m) = f(t, c)$$

Time inputs are available such that

$$(8)\quad g^j(t) = t^j - \hat{t}^j \leqslant 0 \qquad j = 1, \cdots, l.$$

That is, used up resources of service t^j must be at the most equal to the available amount $\hat{t}^j$.

Purchased inputs are restricted by a budget function such that the costs for all c^k must not exceed $\hat{b}$, which gives us the constraint

$$(9)\quad h(c) = pc - \hat{b} \leqslant 0$$

where $p = (p^1, \cdots, p^m)$ is the price vector for the purchased resources $c = (c^1, \cdots, c^m)$. In addition we have the non-negativity constraints

$$t \geqslant 0, \quad c \geqslant 0$$

Necessary conditions for the problem to have a solution can be derived from the Langrangian function:[1]

$$(10)\quad L = f(t, c) - \lambda^j g^j(t) - \mu h(c), \quad j = 1, \cdots, l$$

The general procedure is to take the partial derivative of L with respect to one variable at a time and set the resulting expression equal to zero. We thereby arrive at a set of equations, giving us the necessary conditions for $0 = f(t, c)$ to be a maximum under the given constraints. In addition, some other conditions must be fulfilled as specified below. The following conditions are of interest here: Either $t^j = 0$ or

$$(11)\quad \frac{\partial L}{\partial t^j} = \frac{\partial f}{\partial t^j} - \lambda^j \frac{\partial g^j}{\partial t^j} = \frac{\partial f}{\partial t^j} - \lambda^j = 0 \quad j = 1, \cdots, l$$

and either $c^k = 0$ or

---

[1] We apply here the so called Kuhn-Tucker conditions. For an exposition of the theory, see, for example: G. Hadley, Nonlinear and Dynamic Programming (Reading, Massachusetts: Addison-Wesley Publishing Co., Inc., 1964), pp. 185-211.

$$(12)\quad \frac{\partial L}{\partial c^k} = \frac{\partial f}{\partial c^k} - \mu \frac{\partial h}{\partial c^k} = \frac{\partial f}{\partial c^k} - \mu p^k = 0 \qquad k = 1, \cdots, m$$

In addition the following must be true

$$(13)\quad g^j(t) < 0 \Rightarrow \lambda^j = 0$$

$$(14)\quad h(c) < 0 \Rightarrow \mu = 0$$

Necessary conditions are, however, in this case also sufficient, since the objective function (7) $0 = f(t, c)$ is concave and all the constraints linear. By using the conditions

$$\frac{\partial L}{\partial \lambda^j} = 0,\ j = 1, \cdots, l, \text{ and } \frac{\partial L}{\partial \mu} = 0$$

we can actually solve for all variables t and c, and the Lagrangian multiplyers λ and μ.

From (11) we derive $\lambda^j = \frac{\partial f}{\partial t^j}$ for all $t^j \neq 0$, $j = 1, \cdots, l$, and interpret λ^j as the marginal product (in steps per time unit) of time input j. Hence, λ^1, which in some cases may be the only λ^j, is the marginal product of this particular student's time, in other words <u>his learning rate</u> at the margin or simply learning rate if amount of learning is a linear function of time.

From (12) we have that for all $c^k \neq 0$, $\mu = \frac{\partial f / \partial c^k}{p^k}$ which is

the marginal product of purchased input k per money unit. We may call this the marginal product of money, and it must evidently be equal in all uses, that is, in all uses of the given budget $\hat{C}$. It is now possible to derive the wellknown condition

$$(15)\quad \frac{\partial f/\partial c^i}{\partial f/\partial c^j} = \frac{p^i}{p^j},$$

The ratio of the marginal products of any two inputs, i and j, should be equal to their price ratios.

$$(13)\quad g^j(t) < 0 \Rightarrow \lambda^j = 0$$

states that if an excess supply of time inputs exists they are used up to the point where the marginal product is zero. If that happens to the student's time input, the reasons may be that he is facing a limit set by his ability. This, of course, indicates a social misuse of student time. If some other time input turns out to have zero marginal product, there is an excess supply of that resource, which has to be removed in the long run in order to avoid a social misuse of resources. Generally of course such a misuse will be present long before the marginal product is zero. In the long run one must move in the direction of

$$\frac{\partial f/\partial t^i}{\partial f/\partial t^j} = \frac{p^i}{p^j}$$

where p^i/p^j is the ratio of social opportunity costs for the two time inputs i and j, and the easiest way of securing this is simply

to move these inputs over to the c-vector, if one seeks a long run solution. (This does not mean, however, that t^j's are freely movable once the problem is set up in one form or antoher. Also the budget b will differ for each setup).

If money is supplied in excess so that $\lambda = 0$, money will be used up to the point where the last dollar no longer yields any extra output. Somehow this seems unlikely to occur in real life.

We may now ask, what would be the shadow price of the time inputs implied by their availabilities, in other words, which prices would make

$$(15)\quad \frac{\partial f/\partial c^i}{\partial f/\partial c^j} = \frac{p^i}{p^j}$$

true also for the resources that have no market price? There exists, following the above analysis, at least one such resource, namely student time.

Let p^1 be the shadow price of student time; from (15) we then have

$$(16)\quad p^1 = \frac{\partial f/\partial t^1}{\partial f/\partial c^k} \cdot p^k = \frac{\partial f/\partial t^1}{\lambda} = \frac{\lambda^1}{\lambda}, \text{ for all}$$

$$c^k \neq 0,\ k = 1, \cdots, m$$

which is the ratio of the MP of student time to MP of money. Similarly some other time input t^j would be shadow-priced

$$(16')\quad p^j = \frac{\partial f/\partial t^j}{\lambda} = \frac{\lambda^j}{\lambda},\quad j = 2, \cdots, 1$$

So far we have only one student in the system, and the implications are quite straightforward; the optimal allocation of the available

resources can be read off from a system of equations and a shadow price of student time may be calculated, but the results are of no significance, other than maybe for descriptive purposes.

We are now going to generalize to n students, and also make possible a differential use of resources by the students, that is, it will be possible to let some students take more of the resources than others, if that should increase product. First of all the objective function has to be rewritten

$$(17)\quad \text{Max} \sum_{i=1}^{n} O_i = \sum_{i=1}^{n} f_i(t, c) = F(t, c) \qquad i = 1, \cdots, n$$

where the summation is performed over n students. Note that the subscripts, now to be introduced, refer to individuals.

Since student time is a resource that cannot be distributed in more than one way—each student is using up his own student time—we have n constraints

$$(18)\quad g_i^1(t^1) = t_i^1 - \hat{t}_i^1 \leqslant 0 \qquad i = 1, \cdots, n$$

We disregard here, obviously, the possibility of exchanging one student's time for another's, which could be done by using some of the students as teachers. Usually $t_i = t_j$ for all i and j, implying that all students get the same amount of time per production period. However, nothing will change in principle, if students are given differing amounts of time at their disposal, that is, optimum conditions could be arrived at also from individual "time budgets."

We write the equivalent to (8) and (9) as

$$(19)\ \sum_{i=1}^{n} g_i^j\,(t_i^j) = \sum_{i=1}^{n} t_i^j - T^j \leqslant 0, \quad j = 2, \cdots, l$$

$$(20)\ \sum_{k=1}^{m}\ \sum_{i=1}^{n} h_i^k\,(c_i^k) = \sum_{k=1}^{m}\ \sum_{i=1}^{n} p^k\ c_i^k - B \leqslant 0$$

T is the vector of time resources (teacher time, space time, and so on) for all n students (excepting t^1, their own time) and B is the total budget for purchaseable resources.

The procedure we now have to follow is analogous to the one described above. We have a Lagrangian expression

$$(21)\ L = F\,(t, c) - \lambda_i^j\ g_i^j\,(t_i^j) - \mu \sum_{k=1}^{m}\ \sum_{i=1}^{n} h_i^k\,(c_i^k),\ j = 1, \cdots, l$$

all partial derivates of which should be set equal to zero, except for the possibility of a solution where one or more variables are zero. Starting with the time of student i we get (for $t_i^1 \neq 0$)[1]

$$(22)\ \frac{\partial L}{\partial t_i^1} = \frac{\partial F}{\partial t_i^1} - \lambda_i^1 = 0 \quad i = 1, \cdots, n$$

and derive the marginal product or marginal learning rate of the i^{th} student

---

[1] That t_i^1 in an optimal solution would come out as zero is highly unlikely, however, since it would be equivalent to saying that student i cannot learn anything at all, or, more precisely, cannot learn even the first step in the sequence.

$$\lambda_i^1 = \frac{\partial F}{\partial t_i^1} \qquad i = 1, \cdots, n$$

We also have (for $t_i^j \neq 0$)

$$(23)\ \frac{\partial L}{\partial t_i^j} = \frac{\partial F}{\partial t_i^j} - \lambda_i^j = 0 \qquad \begin{matrix} j = 2, \cdots, l \\ i = 1, \cdots, n \end{matrix}$$

hence, marginal product of some time input used by student i other than student time

$$\lambda_i^j = \frac{\partial F}{\partial t_i^j}$$

Similarly (for $c_i^k \neq 0$)

$$(24)\ \frac{\partial L}{\partial c_i^k} = \frac{\partial F}{\partial c_i^k} - \mu \cdot p^h = 0 \qquad \begin{matrix} k = 1, \cdots, m \\ i = 1, \cdots, n \end{matrix}$$

Marginal product of money, then, would be

$$\mu = \frac{\partial F}{\partial c_i^k \cdot p^k} \qquad \begin{matrix} k = 1, \cdots, m \\ i = 1, \cdots, n \end{matrix}$$

We also have the conditions, equivalent to (13) and (14) stating that if there exists excess supply of any inputs, those will be used up to the point where MP = 0.

What do the optimal conditions tell us in this case? For the purchaseable resources it must hold that (for all $c_i^k \neq 0$)

$$(25)\quad \frac{\partial F/\partial c_i^h}{\partial F/\partial c_i^k} = \frac{p^h}{p^k} \quad \text{for all h, k and i}$$

The ratio of the marginal products in other words, of any two purchaseable resources must be equal to the ratio of their prices, and this must hold for each student in the program, assuming perfect distributility among students of all inputs other than student time. If h = k (25) reduces to

$$\frac{\partial F/\partial c_k}{\partial F/\partial c_i} = 1 \Leftrightarrow \frac{\partial F}{\partial c_i} = \frac{\partial F}{\partial c_k}$$

for any two students i and k; in other words, the MP of any purchaseable input must be equal for all students.

If we consider two time inputs (other than student time), say time inputs h and k, we have

$$(26)\quad \frac{\partial F/\partial t_i^h}{\partial F/\partial t_i^k} = \frac{p^h}{p^k}, \quad i = 1, \cdots, n$$

where p^h and p^k are shadow prices that can be uniquely determined by the optimal solution, hence by the availabilities of the resources. If we are contemplating only one such time input, say teacher time, it should be distributed so that

$$\frac{\partial F}{\partial t_j^{teach}} = \frac{\partial F}{\partial t_k^{teach}} \quad \text{for any two students j and k.}$$

In other words the marginal productivity of teacher time should be equal in all uses, that is for all students.

Summing up, the MP_i of any one resource, except student time should be the same in all uses over all students in order to obtain an optimal output for a given number of students. This is what might be intuitively expected, since given that for some resources $MP_j > MP_k$ for two students j and k, it should be possible to increase product by reallocating the resource away from student k, thereby increasing MP_k, and towards student j, with a resulting decrease in MP_j, a process which would be possible until $MP_j = MP_k$. We also have that ratios of marginal products should be equal to the ratios of their prices.

We finally compute the shadow price of student time as follows:

$$(27)\quad p_i^l = \frac{\partial F / \partial t_i^l}{\partial F / \partial c_i^k} \cdot p^k = \frac{\lambda_i^l}{\lambda} \qquad \begin{matrix} i = 1, \cdots, n \\ k = 1, \cdots, m \end{matrix}$$

Since λ^l generally varies over the students, whereas λ in an optimal solution is the same for all students, it can be seen that the shadow price implied by the availability of resources could be expected to be somewhat different for each student. Furthermore, a student's MP with respect to his own time, as pointed out earlier, may be identified as his learning rate at the margin, so that if one students' marginal learning rate is α times another one's the shadow price of his time will also be α times as high, or in terms of time to completion, if one student takes $\frac{1}{\alpha}$ times as long to complete his marginal task, his time will have to be shadowpriced at α times that of the student with whom he is compared. Shadow price then is proportional to learning rate at the margin and inversely proportional

to time to completion of the marginal task. If learning is a linear function of time as suggested by Carroll, then learning rate for a given student is the same over the whole sequence, and MP is simply equal to that rate of learning.

Since time to completion of a learning task usually varies in such a way that the top five per cent of the students learn about five times as fast as the low five per cent, the shadow prices of their time, assuming stable learning rates, should also have the proportions of five to one, the fast student's time being five times as valuable as the slow student's. This seems to make sense intuitively. It must be pointed out that learning rate, however, is determined in the system by the amount of resources devoted to each student. It may very well turn out that an optimal solution to the programming problem will either reduce or increase the range of time to completion. The average productivity of the fast learner is probably higher than that of the slow learner with respect to any of the inputs, and certainly with respect to his own time. It is, however, the marginal productivities that determine resource allocation, and for that reason it may very well turn out that an optimal solution to the programming problem would allocate resources more in favor of slow learners than is being done in schools today.

It may finally be pointed out that the solution presented above, if all variables are different from zero and no excess supply exists, simply is a "general equilibrium" solution, except that student time is restricted to each particular student, thus allowing marginal products of student time to vary over student.

Summary

In this chapter the analysis was extended to incorporate sequences of steps related in such a way that earlier steps are required for later ones. Time to completion in later steps, it was concluded, depends partly on the allocation of resources on earlier steps. This was formalized in a present value cost model. A programming model was developed next, the major features of which was

1) Derivation of optimal conditions for maximizing output over a sequence when students and their entry characteristics are given.
2) Identification of MP of student time with learning rate at the margin.
3) Shadow pricing of resources that have no market price, especially student time.

The relation between the shadow price of student time and learning rate was examined, and the shadow price of a particular student's time turned out to be proportional to his learning rate at the margin.

CHAPTER VII

A STUDY OF EMPIRICAL COST CURVES IN TIME DIFFERENTIATED EDUCATION

In this chapter some empirical findings will be reported using time to completion data taken in a non-experimental learning sequence in mathematics. The student were allowed to take the time they needed to complete a step before going on to the next level. Söderkullaskolan in Malmö, Sweden, was chosen because this school has a comparatively long experience in teaching of this kind. Schools in Sweden are required to individualize instruction to meet the needs of the students; time differentiation, however, has normally not been used for this purpose until very recently.

Data on time required for learning a sequence of mathematics tasks, on mathematics ability, intelligence, marks in school subjects, and family background were taken for all students in grade 7 (13 year olds mostly), this is the first year of the upper department of the Swedish comprehensive nine year elementary school.[1]

The purpose of the empirical study was twofold:

1) To estimate empirical MC curves in order to illustrate the basic cost model developed in Chapter III.

2) To throw some light on the possible determinants of students' time to completion.

---

[1]The most radical change for the students in going from grade six to grade seven is perhaps the alteration from class teachers to subject matter teachers.

The first task is quite straightforward, involving nothing more complicated than the plugging in of economic data concerning teacher cost, costs of books and materials, and so on—even though, as might be expected, there were some difficulties in arriving at accurate cost figures for some of the items and especially for the opportunity cost of students' time.

Some of the most powerful of the determinants of individual variation in time spent in learning evidently were not captured by the data collected, but such determination was not the main purpose of this study. The time distributions are important in themselves, regardless of what places individuals at various positions in those distributions.

The Students and the Data

Since our sample consists of all students on one grade level, the seventh, in a single school, and therefore is not a random sample, it will be necessary to consider the possibilities of bias in the selection as compared to the population of all seventh graders in Sweden.

There were 171 students in the sample, divided up in six classes, 86 were boys and 85 girls. Most of them (155 or 90.6 per cent) were 13 years old in 1969; this is the normal age in Sweden for 7th graders, the school starting age being seven. Thirteen (7.6 per cent) were 12; in exceptional cases students who by a test and on other grounds are judged to be mature enough may start school at the age of six. The three remaining students were over-aged one year, hence 14 years old in 1969, which may reflect either grade repetition or delayed school

entry. Some students are permitted to start school at eight instead of at seven, because of poor health or for some other reason. This age mix is normal for the grade level.

It is usually believed that schools in Sweden are quite similar in terms of characteristics of their student bodies. Only recently has this assumption of similarity between schools been challenged. In a study (1969) of school segregation in the Malmö school system Swedner and Edstrand found that because of a certain segregation in residential areas, there does in fact also exist a social segregation of schools to the effect that some school districts have disproportionately large enrolments of students from either "high" or "low" social background.[1] The neighborhood rule is with rare exceptions a firmly established enrolment principle for elementary schools in Sweden, so it may be seen that some social imbalance in school enrolments could easily result from residential segregation.

Swedner's and Edstrand's investigation is of interest for the present study because Söderkullaskolan was among the schools investigated; all elementary schools in Malmö were in fact included. Using four categories of father's occupation, the same that have been used in this study, they calculated a segregation number for each elementary school in Malmö. This index was computed by the formula:

$$A = \frac{4a + 3b + 2c + d}{a + b + c + d}$$

---

[1]Harald Swedner and Gisela Edstrand, "Skolsegregation i Malmö," (School Segregation in Malmö), mimeographed (Lund: Department of Sociology, University of Lund, 1969).

where a is number of pupils with father (or guardian) in the "highest" social group, b is the number of pupils with father from the next to the highest, and so on. Theoretically, therefore, the quotient could vary from four (all in the highest category) to one (all in the lowest category): actually the range goes from 3.05 to 1.30. With an index of 1.74 Söderkullaskolan ranks in the middle of the elementary schools in Malmö; 36 schools have a higher index and 24 have a lower. By using the data of the present study an index of 1.85 was computed.[1] The expected value for Malmö computed from Svedner's whole sample would be 1.93,[2] implying that Söderkullaskolan has some overrepresentation of students from "lower" social strata in the Swedner-Edstrand sample; our sample has even less of that bias. This would be expected knowing the type of residential area from which the school recruits its students, being a "mixed" one consisting of rather large multi family dwelling blocks, interspersed with a few small blocks of one family townhouses.

Since the social class grouping for the purpose of the present study used the same classification, it is of interest to summarize briefly the rules that were used. The Central Bureau of Statistics 1952 worked out a scale consisting of three groups: I, II and III.

---

[1]The slightly higher value could - except for a possible year to year variation - depend on the fact that so called special classes were excluded from the present study. (Special classes are designed for low ability students and students with other school problems, not likely to benefit from the teaching in usual classes.)

[2]Ibid., p. 26.

Gustafsson and Swedner at the Department of Sociology, University of Lund, 1962 developed a new scale, mainly by subdividing group II into IIA and IIB, and their scale has since then been frequently used for sociological investigations. The following is a summary from "Skolsegregation i Malmö" (translated into English by the present author).[1]

To social group I belong:

1) Owners and leaders of large corporations
2) High ranking civil servants
3) University graduates employed by private enterprises
4) Self-employed university graduates (e.g., lawyers, physicians)
5) Other university graduates

To social group IIA belong:

1) Civil servants of medium rank (e.g., elementary school teachers, postal functionaries)
2) Employers of private enterprises on a medium level (e.g., accountants, engineers)
3) Owners and leaders of small enterprises

To social group IIB belong:

1) Low ranking civil servants (e.g., petty and non-commissioned officers, postal servants)
2) Low ranking employees of private enterprises (e.g., clerks, traveling salesmen, shop assistents)

---

[1] Ibid., p. 8.

To social group III belong:

1) Low ranking government employees

2) Skilled workers

3) Non-skilled workers.[1]

The distribution of students with respect to father's occupation is displayed in Figure 4 (p. 153). School records were used as a source of information. The missing cases are those who could not be categorized, either for lack of information or because titles were not included in the list used. "Expected values" are the distribution from the Swedner-Edstrand investigation.[2]

A comparison with these expected values gives us the same information as the segregation index, and in addition also tells us that the highest occupation group is underrepresented and that the lowest is slightly underrepresented. Figure 5 (p. 154) gives the distribution of family income before tax for our sample, obtained from the local tax authorities in Malmö regarding the parents' income. For several reasons it does not give a totally accurate picture of a family's disposable income; for example, taxes are included and government transfers are excluded, which gives an underestimate of disposable income towards the lower end of the distribution and overestimate of it towards the higher. Income before tax is, however, the most accurate measure obtainable, and is in Sweden accessible upon request.[3]

---

[1] A complete list of occupations, with the Swedish titles, available from the Department of Sociology, University of Lund.

[2] Ibid., p. 26.

[3] "Offentlighetsprincipen," the publicity principle, is applicable

The bias towards the middle is here more obvious than in the distribution of social groups. This is perhaps what might be expected from the type of residential area we are dealing with, since the very low income people cannot afford the rents in these relatively new houses and the very high income people are more likely to chose their dwellings in other residential areas. It may be of some interest also to note that family income correlates in the expected direction, but not very highly, with occupation ($r = -.32$, significant with $p < .01$).[1]

It is interesting and important to note that the bias towards the middle is very slight indeed, if at all present, when it comes to a measure of pupils' intelligence. Table IV (p. 150) gives the distribution and expected values.[2] Marks in Swedish, English, and Mathematics, and the mean of marks in the orientation subjects also provide some information on the nature of our sample. Teachers in the elementary schools of Sweden use a five point scale (one being the lowest and five the highest grade), and they are expected to distribute marks in

---

here, saying that no information, with explicitly stated exceptions, could be hidden from the public by central and local government officials.

[1] A higher correlation may have been obtained by using logarithms of income. There are also factors depressing income, not showing in social status, for example illness, deceased fathers and divorces.

[2] WIT III was used, a group test of general intelligence, developed by Per Anders Westrin at the Department of Educational Psychology, University of Lund. It consists of four parts: Analogies, Contrasts, Number Combinations, and Puzzle (geometrical figures). The two first parts together yield a measure of verbal intelligence and the third and fourth non verbal intelligence. Split half reliability coefficients for 13 year olds is .93 on total score, .91 for verbal and .85 for non-verbal intelligence. Test-retest reliability is .84. For this and other information on the test see WIT III Manual, (Stockholm: Skandinaviska Testförlaget AB, 1969).

such a way that for all students in the whole country at the same level and taking the same courses an approximately normal distribution will occur. To help teachers approach this goal, standardized tests are given to the students' every third year. If this goal is achieved for the students in Söderkullaskolan, the distribution of marks seems to demonstrate that we are dealing with a sample that does not deviate appreciably from a random one. (Tables 5 and 6, pp. 150-151).

For our purposes, then, using grade seven in one school rather than a random sample of students seems to be preferable, since random sampling would introduce an uncontrolled variation in the treatment conditions of the learning situation. In mathematics the six classes were taught in two large groups, consisting of three classes each, by two teachers and one teaching assistant, the two large groups both having the same teachers and teaching assistant. This fact also, it seems, should contribute to ensuring that treatment conditions were roughly the same for all students involved, at least as far as would be possible in a non-experimental setting.

Summing up, then, the distribution of students with respect to fathers' occupation and parents' income, to their own intelligence, and to marks in school subjects seem to indicate that some tentative generalizations should be possible from our sample.

The Mathematics Materials Used

Having reported the characteristics of the learners, their physical environment, and social background, we now turn to the learning tasks involved.

As mentioned no attempt was made experimentally to control the learning situation; to the contrary, great care was taken not to change in any significant way the teaching and learning situation from what it would presumably have been were no data collected. Since the teaching assistent's job to a large extent consists in keeping track of students' progress through the course, it was natural to ask her to make notes of time taken by every student for each task; only minor changes of her usual routines in class were in fact necessary. Not time in each task could be reported since it is also the teacher assistant's duty to register students' absence.

Very fortunate for our purposes was the fact that the "new math" is now taught in grade seven in Swedish schools. Through the whole sequence selected for our study, the subject was elementary set theory. Therefore, it may be concluded with close to certainty that the students in taking the sequence were confronted with something of which they knew nothing beforehand, and that learning outside school was kept to a minimum. For the first part of the sequence, during which no homework was permitted, learning outside of school was virtually eliminated, since only in very exceptional cases would parents be able to tutor their children in set theory.

The material that was used, called IMU (short for the Swedish for individual mathematics instruction, "individuell matematik-undervisning") has been developed at the Department of Educational and Psychological Research, School of Education, Malmö, commissioned by the Central Board of Education in Sweden, and is being distributed

by Hermods, a correspondence school and a producer of school materials.[1]

Most important for our purposes is the fact that each student is allowed to progress at his own individual speed. Teaching in the form of lecturing is kept at a minimum level and mainly concerned with instruction in the use of the materials; some group instruction and also individual instruction, when called for, are provided by the teachers. The material is partly differentiated with respect to difficulty, so that the most able students can choose more advanced and the less able easier work. The same content is taught to all students, however, although to differing degrees of depth. In addition the students are brought back go the "zero level" at different points of the program, where they learn exactly the same, basic mathematics. The upper department - grade levels seven through nine of the Swedish Comprehensive School - is covered by nine so called modules, with no grade level labels. Should some of the students finish all nine modules before the end of the comprehensive school, they will be provided with materials from the "gymnasium" (senior high school, roughly).

Each module consists of three components: the A component is common for all students, and thus of special interest to us, the B and C components are each divided up, into three versions of different

---

[1] IMU Högstadiet Ett undervisningssystem i matematik utarbetat av Skolöverstyrelsen. (IMU, Upper Department. A Teaching System in Mathematics developed by the Central Board of Education; Malmö: Hermods 1970.) The system consists of a number of booklets, tapes for tape recorders, materials for laboratory work (not for the first sequence, the one used in the present study), and transparents for overhead projection.

levels of difficulty. The students discuss with their teachers which version should be chosen, after having taken a diagnostic test following the A and B component respectively. Figure 6 (p. 155) shows the student flow through a module.

There is no required mastery level in the strict sense; rather, the student is encouraged to reach the mastery level that corresponds to his qualifications. A minimum level exists, however, for each component, which the student must reach before being allowed to go on. The amount of learning over and above this will determine the choice of B and C versions.

For each module there is one booklet for the A component, taken by all students, and one for each of the three B and three C versions. In the first module, the one used in this study, there are, however, only two B versions, B1 and B2-3. Each module also has additional review and correction materials, assigned as needed after each diagnostic test; when time was used on such materials it was added as appropriate to the time used in the A, B and/or C components. Some students were also sent back in the "program," in which cases this time also was added. After having proceeded through the whole module the student takes an evaluative test and goes on to the next module. Hence we have data on time to completion of the first module: time in component A for all the students, and time in each of the two B versions and the three C versions according to the students' assignments after the diagnostic tests.

During the time they studied the A component the students in our sample were not allowed to take their mathematics materials home with them, which contributes towards keeping outside learning for this

component at a minimum. From the B component and on they were required or perhaps rather permitted, to do some homework. The time was reported by the students for the purpose of this study - usually they are not required to do that. Great care was taken in order to make the students understand that this report was not intended for the teachers and that whatever was reported should not be revealed. This procedure, it was believed, should minimize the risk that students used the report for the purpose of impressing their teachers. It is probably best, nevertheless, to look upon these reports of time used at home with some suspicion; there seems to exist no reliable source of information for time spent in learning outside the school. The teachers were not willing to extend the period during which homework was prohibited beyond the A component.

Analysis of Time to Completion Data and Derivation of Empirical Marginal Cost Curves.

In this section we will attempt to describe the time curves and add economic data in order to obtain MC curves. The exercise must be seen as an example of how this can be done. No comparisons are made with other teaching situations—once again no controlled experiment was intended—and the statistical analysis is only preliminary, mainly consisting of curve plotting.

Table 7 (p. 152) gives the statistics for the time distributions over the whole sequence. One student did not complete any of the steps, and another one completed only step A. Looking at the distribution of time for the A step, which gives the most reliable information as far as the sum total of time used is concerned, we can see that there is a

difference of about one to four between the slowest and the fastest student. Not counting the top and bottom two per cent would still make the difference one of about one to three. This is somewhat lower than the difference of one to five or six that is usually reported. As can be seen, however, the difference tends to increase somewhat on the higher levels, in spite of the fact that those levels are differentiated with respect to difficulty. (The slow student tends to take the easy version.)

With the exception of B1, with only ten students, and time in school for C2, all distributions are positively skewed, which means that there is a long tail to the right of the distribution; a small number of students take a very long time for completion.

The cumulative frequency distributions of sum of time in school and at home for all steps are shown in Figures 7 through 12 (pp. 156-161). The positive skewness appears as a comparatively long and flat upper part of the curves. The distribution of C2, however, seems to be approximately normal. In Figure 13 (p. 162) this curve is plotted on a normal probability paper, and comes out as can be seen, rather close to a straight line. The most probable reason for C2 being normal and not skewed is that, by selection, the extreme students at both ends of the total sample had been sorted out; they are in C1 or C3.

It would be premature to conclude from this study, using only one curricular subject and just a few time distributions, that ability as measured by time to completion is positively skewed. Motivational factors may enter into this and will be taken up for discussion later on. There is, however, something in the nature of time as an ability

scale that makes this positive skewness highly plausible. As will be seen, this has important economic implications. When using an arithmetic time scale three hours, for example, is exactly three times as much as one and six is three times two. But when used as a measure of human ability arithmetic equidistance in the units of the scale may be inappropriate. If, for example, one person could run the 100 yards dash in, say, ten seconds and another one in twenty, we would affirm that the first one is twice as fast. We normally, however, would not conclude that he is twice as able an athlete; the difference, intuitively, would seem to be much larger than that. Improving one's capacity from 20 to 19 seconds would for most people be quite easy, whereas improving the world record on the 100 yards dash by even one tenth of a second is an extreme and rare achievement indeed.

A difference of one arithmetic time unit at the upper end of the scale, therefore, seems to carry less weight than a difference of one time unit at the lower end of the scale. In other words when going from the slowest pupils, taking comparatively long times, to ever shorter time, each arithmetic reduction in time entails a progressively greater rate of decrease in physical time units. It is of some interest, therefore, to examine what happens to the skewness of the time distributions when those distributions are scaled logarithmically.

Human ability is often thought of as being normally distributed. Some statistical-theoretical foundation for this can perhaps be found in the "law of the large numbers." That measures of ability, such as IQ and aptitude scales, often do come out very nicely as approximate normal curves, may on the other hand be a construct, since items not

supporting normalcy are sorted out when tests are developed. If we assume, nevertheless, that ability is normally distributed when "correctly" measured, whatever that means, then ability as measured by time to completion would be normally distributed in the logarithms, if indeed the time scale as an ability measure is logarithmic.

To explore this possibility the cumulative time to completion distributions of two steps (A and B2-3 where most of the students were included) were plotted on a semi-log paper (Figure 14, p. 163) and logarithms of the same distribution were plotted on a normal probability paper (Figure 15, p. 164). As can be seen, those log distributions do in fact (by inspection) come close to normalcy in the logarithms. The plots in Figure 11 (p. 160) have, approximately, the shape of normal and non-skewed, cumulative distributions, moreover, the same distributions on probability paper come out as almost straight lines, showing the same thing.

The importance of the cumulative frequency distribution becomes clear when we consider its relation to an MC-curve. Assume a case where costs are perfectly proportional to time in learning. Then the MC function is simply the inverse of the cumulative time frequency distribution with the time scale converted to a cost scale by the appropriate weight. The more complicated cases of MC curves will be transformations of this inverse. Introducing costs of supplies, equal for all students per step (S), would for example simply mean adding a constant to the curve.

Cost curves, as plotted in Figure 16 through 21 (pp. 165-170), were computed from the empirical data. Three different measures of student

time were used: 1) student time valued at zero, 2) student time valued at 5.15 Swedish kronor (1969) and 3) student time valued at 3.00 Swedish kronor per hour. In the second case wage per hour before tax in the age group 14-19 was used, and in the third case the value of student time was assumed to develop linearly, from age at school entry when the value was assumed to be zero to the midpoint of earnings in the age group 14-19.[1] This procedure is less satisfactory than the shadow pricing of the programming model; it treats all students' time as of equal value whereas the shadow price derived in a programming model would be inversely related to the time the individual takes. Shadow prices thus tend to flatten the curve as compared to the use of mean values of foregone earnings.

Other costs were derived from the (ex post) budget of the Malmö school system. Since the breakdown of this budget is rather detailed and the principle is to credit costs to "real" rather than merely "accountant" entries it is believed that figures are fairly accurate. The method is not flawless it must be admitted; school space costs are especially difficult to estimate. The costs can be itemized as follows:

Teachers and other school personnel

Free school materials (books, stationery, and so on)

Equipment

---

[1]Source: Svenska folkets inkomster. Betänkande avgivet av låginkomstutredningen, SOU 1970:34 (The Swedish Income Distribution, Report from the Low Income Committee, Publications from Government Committees, 1970, No. 34), (Stockholm: Inrikesdepartementet, 1970), p. 101.

School Space

Administration

School lunches

School health

Busing

AV materials

Experts (e.g., consultants, educational psychologists, curators)

"Other"

The per student cost was calculated to 3.90 Swedish kronor per period in school and adjusted downwards by subtracting costs of books and supplies (to be replaced by the costs of the IMU material) and the difference between the cost of using a teacher assistant and the cost of a third teacher. The per student cost exclusive of materials and student time thus arrived at was 3.35 Swedish kronor per period in school. As can be seen, even the lower of the two estimates of student time (excluding zero) is almost as high. The IMU material has a market price from which the cost per step was calculated to be 5.00 Swedish kronor for each student.

It can be seen from the graphs that the skewness of the time distribution, as may be expected, has resulted in a sharp rise of the MC curve (approximately for the ten per cent slowest students in this sample). Quite aside from the obvious "social" arguments that can be mobilized for finding ways and means of reducing time differentials among students, there apparently are also purely economic reasons for wanting to do so. Depending upon the costs of reducing the time differentials, this may or may not, however, lower the MC curve. If

the MC-curve is not lowered when time differentials are diminished, then in turn any eventual benefit change has to be taken into account when making decisions concerning measures aimed at reducing the time to completion for the slowest learners.

Relationships among Variables

In order to explore somewhat how the variables in the sample relate to each other, a correlation analysis was used. Table 3 (p. 149) shows the zero order coefficients for time to completion of the first step, IMU A, with final achievement on the whole module; ability in math;[1] verbal, non-verbal and total intelligence;[2] father's or guardian's occupation;[3] family income; marks in selected school subjects in grade 6 (spring semester) and grade 7 (fall semester). The table also displays interrelations among the "independent" variables.

As can be seen we have the usual correlation in the neighborhood of .4 to .7 among measures of school achievement. Remarkable is the absence of a significant correlation between either family income or father's occupation and any of the other measures (with two exceptions where r is quite low and not significant, with $p < .01$). There is a correlation between the two family background variables but smaller than perhaps might be expected.

---

[1] as measured by an instrument developed for the evaluation of the IMU project by the Department of Educational and Psychological Research, School of Education, Malmö.

[2] as measured by WIT III.

[3] as measured by the scale described on pp. 120-122.

Time to completion of IMU A correlates moderately and in the expected direction ($r < 0$) with intelligence and with ability and achievement measures—which is what Kim also found. Again noteworthy is the absence of correlation with social background variables. When time to completion of steps B2-3, C1, C2, and C3 was related to the other measures it turned out, however, that the correlations mostly were near zero.[1] For all versions of the C-step r was never significantly different from zero ($p < .05$) whereas time to completion of B2-3 correlates in the neighborhood of -.2 with most of the other variables, as can be seen from the correlation matrix.

The structure of the correlation matrix raises several questions. Why is it that ability as measured by time to completion of a learning task does not correlate more highly with other ability measures? And why does this already moderate correlation go down as new steps are added. This decline and finally disappearance of a significant correlation may have several possible explanations:

1) measurement errors
2) when students choose between the versions of B and C they are thereby divided up in ability groups to some extent, thus becoming more homogeneous within the chosen groups
3) when homework is introduced, r becomes lower (homework, as we recall, being permitted from step B and on)

---

[1] B1 was not included in the correlation study since only 10 students chose this version.

4) explanation offered by the economic theory of utility maximization in time allocation.

One obvious source of measurement error is the possibility, not to say probability that some students over-perform relative to the criterion; in other words, we were not able to measure exactly when they reached criterion and no more. From the correlation between time to completion of A with the final test ($r = -.46$, see Table 3, p. 149) it seems plausible that the slow learner show the least overperformance. That homework could reduce r may entail (partly at least) an aggravated measurement error, since students' reports on homework probably are quite unreliable. Only 117 out of 171 students reported any homework at all, but this is not conspicuously low, since the policy is to reduce homework to a minimum in Swedish elementary schools. One cannot be sure, however, that none of the other 59 students learned anything at home. Also it is very probable that reports in some cases of time used at home in studying mathematics were guesswork rather than accurate knowledge.

Time at home in both B and C has a low but significant positive correlation with the result on the final test, which perhaps supports the "overperformance" explanation. Time spent at home in C, but not in B, is also (positively and significantly) correlated with marks.

With all these explanations, it is still true, however, that the moderate measure of r for the first, and more controlled, step is in very good agreement with what others have found. Kim's study, for example, was a more carefully controlled experiment in educational measurement (which this study was not intended to be), yet his

correlations are only slightly higher than those reported here for step A. It would be satisfactory to get a more powerful explanation than the above 1) to 3), and we will now turn to economic theory and see if it can provide one.

In the theory of timing in education, as laid out especially in Chapters II and VI of this study, we have treated learning in schools as a production problem using either a cost-minimization or product-maximization approach. Student time has been treated as a societal resource, and we have completely ignored the fact that students may want to use their time in a different way. Following Becker's time allocation theory[1] we may instead treat the students as utility maximizers, restricted by a time budget. This was the approach of Adelman and Parti when developing a theory of student motivation upon which they based a study of student verbal achievement.[2] They viewed the students' use of time in school as directed towards maximizing a stream of present and future rewards, among which may be parental and teacher approval, expected future earnings, peer acceptance, and leisure. For the following analysis we also add out-of-school time and rewards associated with activities for which such time is an input.

Students are assumed to maximize utility functions:

(1) Max U (t) subject to

---

[1]Becker, "A Theory of the Allocation of Time."

[2]Irma Adelman and Michael Parti, "The Determinants of Student Achievement: A Simultaneous Equation Approach" (unpublished manuscript, Chicago: Northwestern University, 1971), pp. 2-8.

$$(2) \quad \sum_{i=1}^{n} t_i - \hat{t} = 0$$

where $t = t_1 \cdots t_n$ is time spent in activities 1 to n, and $\hat{t}$ is available time. We assume for simplicity that physical inputs in those activities are either freely available or not necessary. The student's utility, in other words, depends only on how he spends his time. If free to divide his time between school and non-school activities the student would, therefore, choose the time input combination that gives him maximum satisfaction in very much the same way as a consumer maximizes utility by allocating his given budget on different market goods and services.

Let $t_s = t_1 \cdots t_m$ be school activities, and $t_h = t_{m+1} \cdots t_n$ be non school activities. Then a necessary condition for utility maximization would be

$$(3) \quad \frac{\partial U}{\partial t_i} - \lambda = 0 \qquad i = 1 \cdots n$$

where λ is a Lagrange multiplyer. It follows that

$\lambda = \frac{\partial U}{\partial t_i} = \frac{\partial U}{\partial t_j}$ all i and j (independently of whether they belong to the t_s or the t_h vector). In other words, marginal utility should be the same in all uses of time, whether in-school or out-of-school time.

Usually, however, the student is <u>not</u> free thus to divide his time. He is restricted to use a certain specified amount of time in school, which introduces two new constraints instead of constraint 2) above.

$$(4)\quad \sum_{i=1}^{m} t_{s_i} - \hat{t}_s = 0$$

leaving the rest of total time for non-school activities, which may be expressed:

$$(5)\quad \sum_{j=m+1}^{n} t_{h_j} - \hat{t}_h = 0$$

where t_h is time available for out of school activities. We have now the new necessary conditions, introducing Lagrange multiplyers λ and ρ:

$$(6)\quad \frac{\partial U}{\partial t_{s_i}} - \lambda = 0 \qquad i = 1 \cdots m$$

$$(7)\quad \frac{\partial U}{\partial t_{h_j}} - \rho = 0 \qquad j=m+1 \cdots n$$

When these constraints are introduced, marginal utility may be (probably is), different for time used as inputs in school activities and in non-school activities. We can assert, however, that

$\frac{\partial U}{\partial t_{s_i}} = \frac{\partial U}{\partial t_{s_j}}$ for all i and j, in other words marginal utility will be equal in all uses of <u>school time</u> (and equivalently for marginal utilities of non-school time as inputs in non school activities).

School activities do, however, as was pointed out by Adelman and Parti[1] include also "non-approved" activities. Daydreaming, looking

---

[1] <u>Ibid</u>., pp. 2-4.

out of the window, playing tricks on the teacher, and similar non-academic uses of school time may not per se seem to be very rewarding as we see it, but the proportion of time used by a student in non-approved activities must be seen in relation to the alternatives open to him, which may for him not seem to be very rewarding either.

From this theory we would predict, or at least find very plausible, that students do spend some time in school in pursuit of non-academic goals and that this time proportion will vary among individuals. The empirical time measure in the present study is, therefore, a sum of time inputs used in academic pursuits and time inputs used in non-approved activities. Suppose now that the "true" relation between time to complete a learning task (excluding, that is, all time spent in school in non academic activities) and intelligence is perfect and linear, that is, the "true r" is unity. If proportion of time spent in non academic activities is unrelated to intelligence, we would expect the measured r to be lower than the "true r" because of the introduced spread around the regression line. More likely, perhaps, is the possibility that highly intelligent students, as result of the prevalent reward system in schools, would tend to devote a higher proportion of their time to academic activities. Then non-linearities would depress the measured r, which therefore will tend to go down also in this case. If the proportion of time spent in learning ("true time") as compared to that spent in non-approved activities will tend to be larger for high ability students, we may here have an explanation also of the skewness of the time curves, hence of the sharp upward turn of marginal cost curves. Measured time, in other words, would

tend to depart from time actually used as input in learning in such a way that it will tend to overestimate time spent in learning more and more as we move from fast and able students towards slow low ability students.

If we now introduce homework, it can be seen from equations 6 and 7 that measured r in all probability will departure even more from a "true" relation, the reason being that the time input to be used in homework has to be taken out of a different "time budget" altogether. Let $\frac{\partial U}{\partial t_s}_{math}$ be the MU of time from "study math in school" and $\frac{\partial U}{\partial t_h}_{math}$ MU of time spent studying math at home. In general we will have

$$\frac{\partial U}{\partial t_s}_{math} \neq \frac{\partial U}{\partial t_h}_{math}$$

(with equality only by coincidence, in which case (4) would no longer constrain the student). It can be seen that the student would add time in studying math at home only as long as the MU of doing so is less than the MU of the alternatives open to him. In "equilibrium" one would expect

$$\frac{\partial U}{\partial t_s}_{math} < \frac{\partial U}{\partial t_h}_{math}$$

The alternatives open to the student, when out of school (teasing sisters and brothers, playing softball, eating ice cream, and so on) are much more powerful competitors for the use of his time than alternatives to academic work within school (daydreaming, scribbling in

textbooks, and so forth). Again his decision as to what to do, may or may not relate to his intelligence, but in either case the result of introducing homework probably would be to further depress the r. This is so since time in studying reported by the student would include, in the same way as time spent in school, some time used in non-academic activities. Because of the much more powerful "distractors" at home, this proportion would probably be higher than in school for most students.

At this point one might well wonder whether this complicated economic apparatus is at all necessary or useful in explaining students' use of time, in school or at home. Why not simply say that students are more or less motivated, for example, to study mathematics. This is, however, to overlook a very significant part of the economic argument. Besides, there seems to be some research evidence to the effect that student interest is only a weak determinant of learning.[1] The economic point is that the amount of time a student is willing to spend on mathematics is dependent not only on his subjective evaluation of the rewards of doing so, but also on his simultaneous evaluation of the alternatives open to him. According to this theory, a student may well answer a questionnaire in such a way as to indicate a very high interest in mathematics; yet he may not be willing to spend very much time studying it, since he has an even higher interest in getting peer approval by trying to bring the teacher close to a nervous breakdown (or in less obvious and therefore less noticed ways).

---

[1] Carroll, "Problems of Measurement," in Mastery Learning, ed. by Block, p. 32.

Summing up then, it would seem highly likely that a hypothesized relation between time in learning as an ability measure and other ability measures would be clouded as long as we cannot measure the exact amount of time used as an input in learning. As can be seen this is in accord with Carroll's treatment of time as "time needed" and "time spent" in learning, which in the above analysis would be equivalent to time actually used as an input in learning and time spent in school, regardless of its use. Hopefully, the economic analysis could also contribute by throwing some light on what would be involved in trying to make time spent in school more nearly like time needed "under ideal conditions,"—ideal, that is, in maximizing the efficiency of in-school time with respect to the production of learning. The above mentioned study by Adelman and Parti was addressed to exactly this problem.

Predictability of Time to Completion Patterns

For the purpose of educational decision making there are two problems involved in predicting time to completion of learning tasks:

1) We wish to know whether observed data will be approximately repeated under similar conditions, and
2) We wish to know, in addition to the above, approximately where on the MC curves a particular student (or perhaps rather a particular type of students) will be located.

Knowing that time-to-completion patterns are largely repetitive for successive age cohorts would be a sufficient basis for many educational decisions. Economic evaluation of teaching methods, where

we are faced with an "all-or-nothing" problem, is an example; in other words when only one of the evaluated methods may be chosen, it will not be necessary to predict individual students' learning rate. Determining student flow through a system would also be an example, if a "proper" self selection by the students could be assumed. If we want to target certain measures to a particular group of students within the distribution, say the 10 per cent slowest, without being especially interested in who those students are, it would also be enough to know that a particular time curve could be expected approximately to repeat itself.

Although the characteristics of the student body in the present investigation are such as to make it very plausible that the time patterns obtained are not merely coincidental, we of course have no basis for concluding anything about the stability of the time pattern from the present study alone. The negative correlation between time to completion and ability (as well as achievement measures) has been shown in several studies, however, and should be regarded as established, together with the fact that this correlation is rather moderate. Repeated studies are necessary in order to investigate further the shape of the time distributions, to see, for example, if the skewness is a characteristic typical of such distributions.

Turning now to the second prediction problem, queries involving allocation of "marginal" students within a system or between systems would require an approximate knowledge of who those students are. From the last section of this chapter and from other research it seems that the basis for such predictions is not very good. To illustrate the

point a plot was made of the cumulative frequency distribution of time in IMU A and of the means for different intelligence groups (Figure 22, p. 171). We can now compare what would have been the result of "prediction" of the time to completion distribution by perfect knowledge and of using intelligence as a prediction instrument. As can be seen, the agreement is not overwhelming.

Summary

In this chapter some empirical findings were presented. A sequence of learning tasks in mathematics, taken by students in the upper department (grade 7) of a Swedish elementary school, was investigated with respect to time to completion and these measures were used for constructing MC cost curves for different steps in the sequence. The skewed time distribution was shown to result in sharply rising MC curves for the slowest ten per cent of the students.

Time in learning was related to intelligence, ability and achievement measures and time in step A showed a moderate, negative correlation with these, whereas the correlation for later steps got closer to zero and finally disappeared. A tentative explanation was offered by applying an Adelman-Parti specification in time allocation theory. Students were said to maximize utility by allocating some time to studying mathematics and some to non approved activities, measured time in school being a sum of time in those different uses.

Predictability was said to be low as far as individual time to completion goes. Not all decision making, however, requires such predictability—only a reasonable assurance of approximate stability

in time patterns from year to year for successive age cohorts would be necessary. Research has yet to establish the existence of such stability.

APPENDIX

TABLES AND ILLUSTRATIONS TO CHAPTER VII

TABLE 3

INTERCORRELATIONS AMONG VARIABLES

(zero order correlation matrix)

		1	2	3	4	5	6	7
1. Time to completion, IMU A			.28	(.07)	(.02)	-.46	-.53	-.29
	N.		159	117	116	169	170	162
2. Time to completion, IMU B2-3				.26	.21	-.14*	-.13*	(-.02)
	N.			117	116	159	159	159
3. IMU B Time at home					.52	(.11)	(.04)	(-.01)
	N.				116	117	117	112
4. IMU C Time at home						.21*	(.13)	(.10)
	N.					116	116	111
5. Final test, first modul							.68	.61
	N.						169	162
6. Ability in math, IMU								.54
	N.							162
7. Intelligence verbal								
	N.							
8. Intelligence non-verbal								
	N.							
9. Intelligence total								
	N.							
10. Father's occupation								
	N.							
11. Family income								
	N.							
Marks Grade Level 6								
12. Swedish								
	N.							
13. Mathematics	N.							
14. English								
	N.							
15. Average of orientation subjects								
	N.							
Marks Grade Level 7								
16. Swedish								
	N.							
17. Mathematics, general course								
	N.							
18. Mathematics, special course								
	N.							
19. English, general course								
	N.							
20. English, special course								
	N.							
21. Average of orientation subjects								
	N.							

( ) Coefficients within parenthesis are not significantly different from zero with $p < .05$.

*Starred coefficients significantly different from zero with $p < .05$.
All other coefficients are significantly different from zero with $p < .01$.

8	9	10	11	12	13	14	15	16	17	18	19	20	21
-.43 162	-.40 162	(.11) 164	(.00) 170	-.37 170	-.50 170	-.34 170	-.36 170	-.35 170	-.44 45	-.37 125	(-.03) 26	-.39 143	-.46 170
(.03) 154	(.02) 154	(-.06) 155	(.00) 159	-.15 159	-.18 159	(-.06) 159	-.14* 159	-.16* 159	-.29* 35	(-.07) 124	-.40* 20	(-.03) 139	(-.19) 159
(.02) 112	(.03) 112	(.04) 114	(-.04) 117	(-.02) 117	(-.02) 117	(.09) 117	(.00) 117	(.12) 117	(.20) 34	(.04) 83	(.15) 21	(.02) 96	(.05) 117
(.13) 111	.18* 111	(.11) 112	(-.06) 116	(.10) 116	.19* 116	.27 116	(.14) 116	.29 116	(.18) 36	.25* 80	.58 19	.18* 97	.25 116
.58 162	.67 162	(-.06) 164	(-.06) 169	.60 169	.77 169	.60 169	.65 169	.59 169	.72 44	.91 125	(.26) 27	.60 143	.70 169
.58 162	.64 162	(-.06) 165	(.09) 171	.60 171	.68 171	.53 171	.54 171	.52 171	.39 46	.55 125	(.26) 27	.50 143	.57 169
.43 162	.83 162	(-.07) 157	(.07) 162	.64 162	.52 162	.62 162	.56 162	.58 162	(.25) 40	.47 122	(.30) 24	.58 138	.53 162
	.82 162	(-.10) 157	(.09) 162	.38 162	.50 162	.41 162	.37 162	.36 162	.41 40	.48 122	(.20) 24	.41 138	.47 162
		(-.09) 157	(.10) 162	.60 162	.58 162	.61 162	.56 162	.54 162	.46 40	.54 122	(.25) 24	.61 138	.59 162
			.32 165	(-.06) 165	-.14* 165	(-.10) 165	(-.10) 165	(-.08) 165	(-.12) 44	(-.08) 121	(-.04) 26	(-.05) 139	(-.12) 165
				(.11) 171	(.08) 171	.16* 171	(.06) 171	(.06) 171	(.03) 46	(.10) 125	(-.04) 27	(.06) 143	(.05) 171
					.59 171	.74 171	.73 171	.75 171	.48 46	.47 125	(.24) 27	.75 143	.66 171
						.60 171	.70 171	.54 171	.55 46	.78 125	(.16) 27	.53 143	.67 171
							.64 171	.65 171	.45 46	.42 125	.68 27	.75 143	.60 171
								.65 171	.43 46	.58 125	(-.05) 27	.68 143	.74 171
									.42 46	.45 125	.34* 27	.70 143	.71 171
										0	(.28) 24	.47* 21	.63 46
											(.50) 3	.44 122	.61 125
												0	.39* 27
													.67 143

TABLE 4

PERCENTAGE OF STUDENTS IN EACH INTELLIGENCE GROUP

N=162

Stanine points	1	2	3	4	5	6	7	8	9	Total	Mean stanine
Verbal Part											
Sample	.6	4.9	13.6	17.9	12.3	22.8	17.9	6.8	3.1	100	5.3
Expected	4.0	7.0	12.0	17.0	20.0	17.0	12.0	7.0	4.0	100	5.0
Non Verbal Part											
Sample	6.2	7.4	15.4	21.0	16.7	17.9	12.3	2.5	0.6	100	4.5
Expected	4.0	7.0	12.0	17.0	20.0	17.0	12.0	7.0	4.0	100	5.0
Total											
Sample	3.1	6.8	11.1	18.5	22.2	21.6	11.7	3.7	1.2	100	4.9
Expected	4.0	7.0	12.0	17.0	22.0	17.0	12.0	7.0	4.0	100	5.0

Source of expected values: WIT III Manual

TABLE 5

PERCENTAGE DISTRIBUTION OF MARKS AT THE END OF THE SPRING SEMESTER, GRADE LEVEL 6

N=171

Marks	1	2	3	4	5	Total	Mean
Swedish	2.9	20.5	42.7	28.7	5.3	100	3.1
Mathematics	5.8	23.4	45.0	20.5	5.3	100	3.0
English	1.2	21.6	39.2	31.6	6.4	100	3.2
Expected	7.0	24.0	38.0	24.0	7.0	100	3.0

Source for expected values: Läroplan för grundskolan (Curriculum for Comprehensive School), p. 90.

TABLE 6

PERCENTAGE DISTRIBUTION OF MARKS AT THE END OF THE FALL SEMESTER, GRADE 7.

N=171

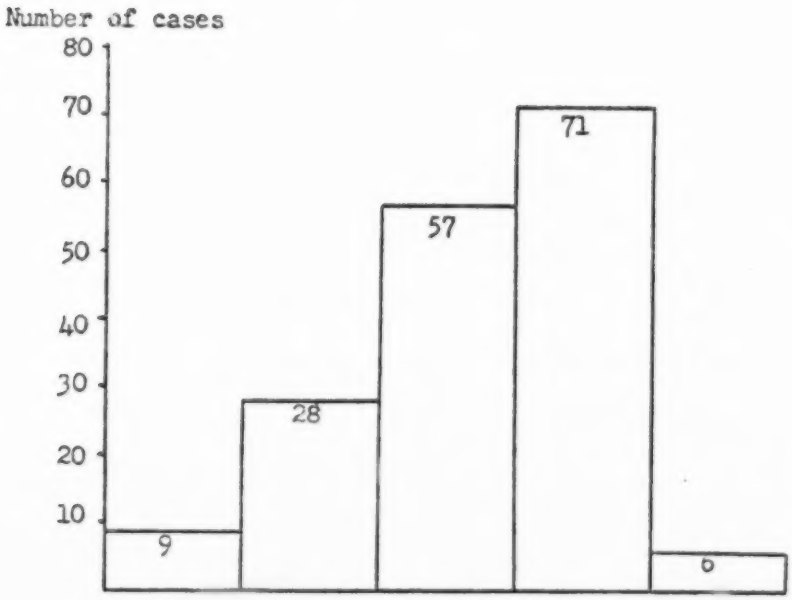
	1	2	3	4	5	Total	Mean
Swedish	1.8	26.9	45.6	23.4	2.3	100	3.0
Math G (N=46)	28.3	23.9	34.8	10.9	2.2	100	2.3
Math S (N=125)	7.2	24.0	37.6	25.6	5.6	100	3.0
English G (N=27)	7.4	25.9	59.3	7.4	0	100	2.7
English S (N=143)	7.0	21.0	42.7	23.8	5.6	100	3.0
Expected	7.0	24.0	38.0	24.0	7.0	100	3.0

In mathematics and in English students study either a "general" course with a more practical orientation, G in the table, or a "special" course, which is more theoretical, S in the table. (With the introduction of the IMU material, see pp. 120-124, this division has in math become a mere formality.) Source of expected values: Läroplan för grundskolan (Curriculum for the Comprehensive School), p. 90.

TABLE 7

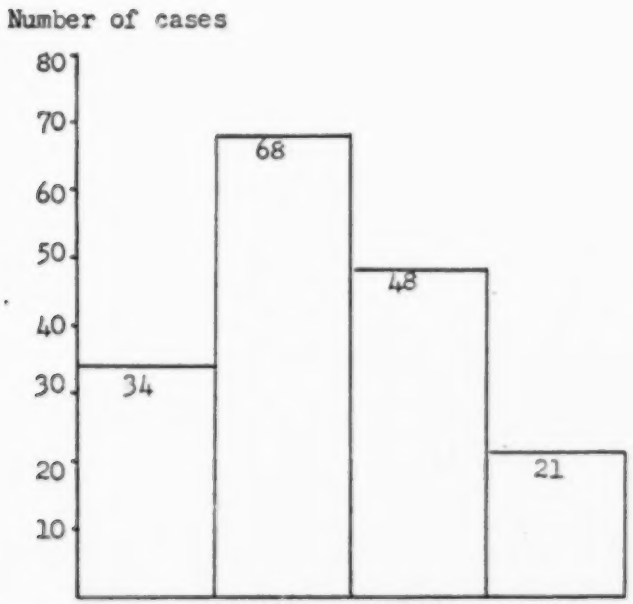
DISTRIBUTION OF TIME FOR COMPLETION OF LEARNING TASKS IN MATHEMATICS. TIME IN MINUTES.

	N	Minimum	Mean	Maximum	Standard Deviation	Skewness
Time in school						
A	170	240	532	1040	141	.7
B1	10	320	624	880	168	-.3
B2-3	159	240	629	1440	173	.9
C1	16	320	672	2000	449	1.8
C2	47	160	374	640	118	.0
C3	106	240	434	960	135	1.0
Time at home						
B	117	8	104	535	99	2.0
C	116	5	106	340	79	.9
Sum of time						
B1	10	328	629	915	173	-.8
B2-3	159	375	701	1515	169	1.1
C1	16	320	712	2075	461	1.9
C2	47	442	722	109	109	.1
C3	106	240	514	960	140	.7



Social Group	I	IIA	IIB	III	Missing
Per cent	5.3	16.4	33.3	41.5	3.5
Adjusted Per Cent	5.5	17.0	34.5	43.0	—
Expected Per Cent	11.5	16.3	26.7	45.5	—

Figure 4.—Number of pupils with fathers in occupation groups I to III. Expected per cent from Swedner and Edstrand "Skolsegregation i Malmö," p. 26, showing the distribution for their whole sample of all elementary school children in Malmö 1968 (grades 1 through 9).



Per cent	20	40	28	12
Expected Per cent	27	30	21.5	21.5

Figure 5.—Family income distribution. Expected values computed from a sample by "Låginkomst-utredningen" (Low Income Committee), "Svenska folkets inkomster," SOU 1970:34 (The Swedish Income Distribution, Report from the Low Income Committee, Publication from Government Committees, 1970 No. 34) p. 221. Values in 1969 Swedish kronor.

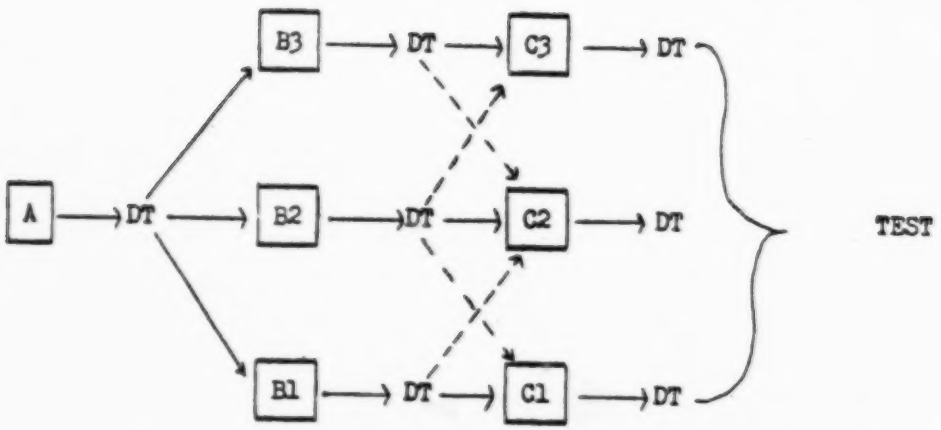


Figure 6.—Student flow through a module of the IMU material. DT = Diagnostic test. The whole module is completed by taking an evaluative test.

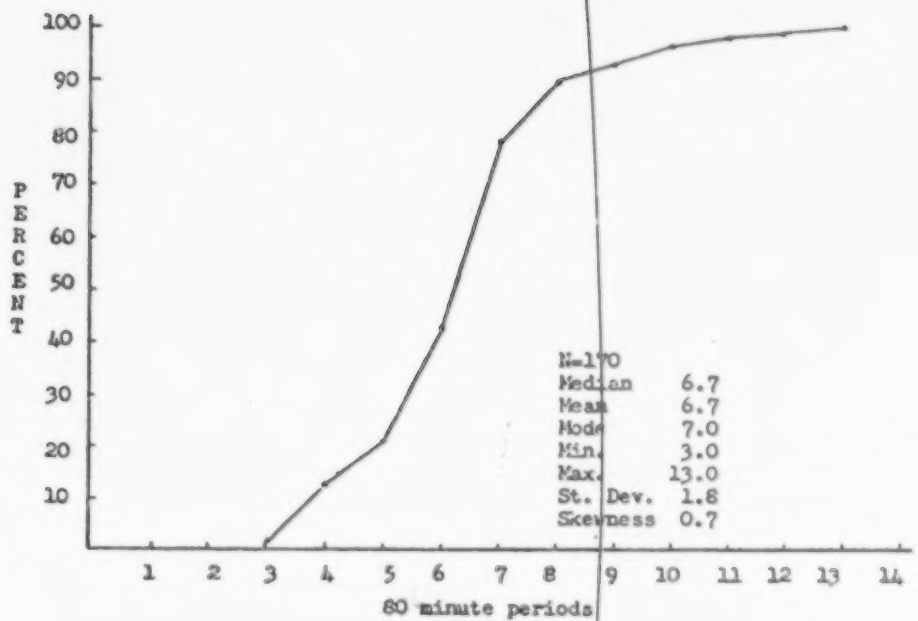


Figure 7.—Cumulative frequency distribution of time to completion, IMU A.

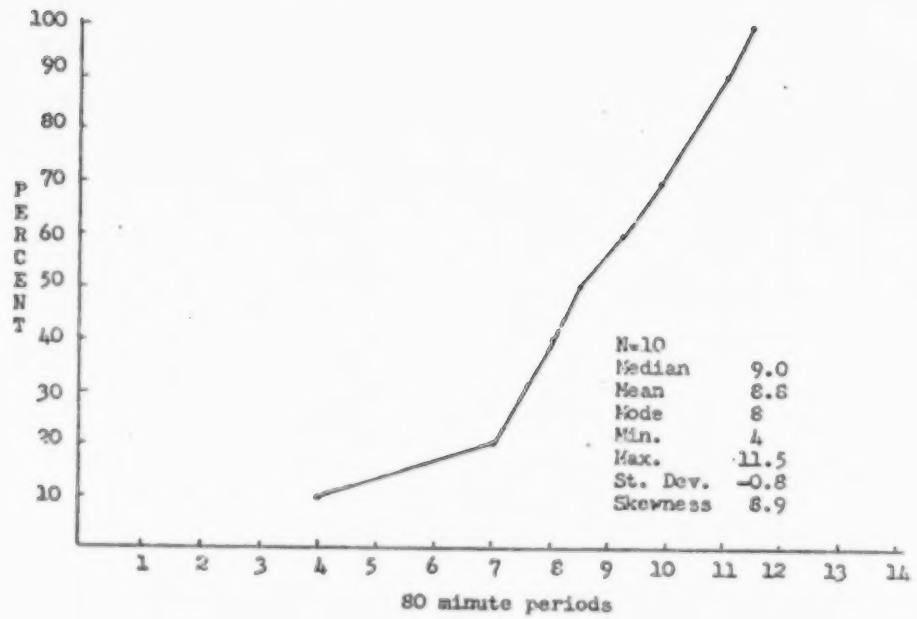


Figure 8.—Cumulative frequency distribution of time to completion, IMU B1.

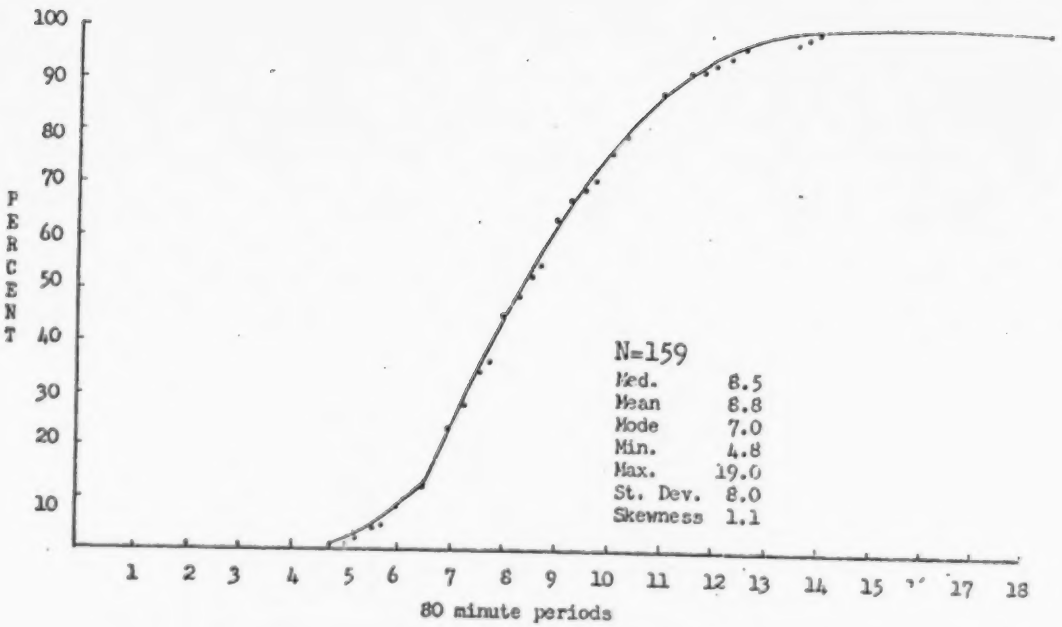


Figure 9.—Cumulative frequency distribution of time to completion, IMU B2-3.

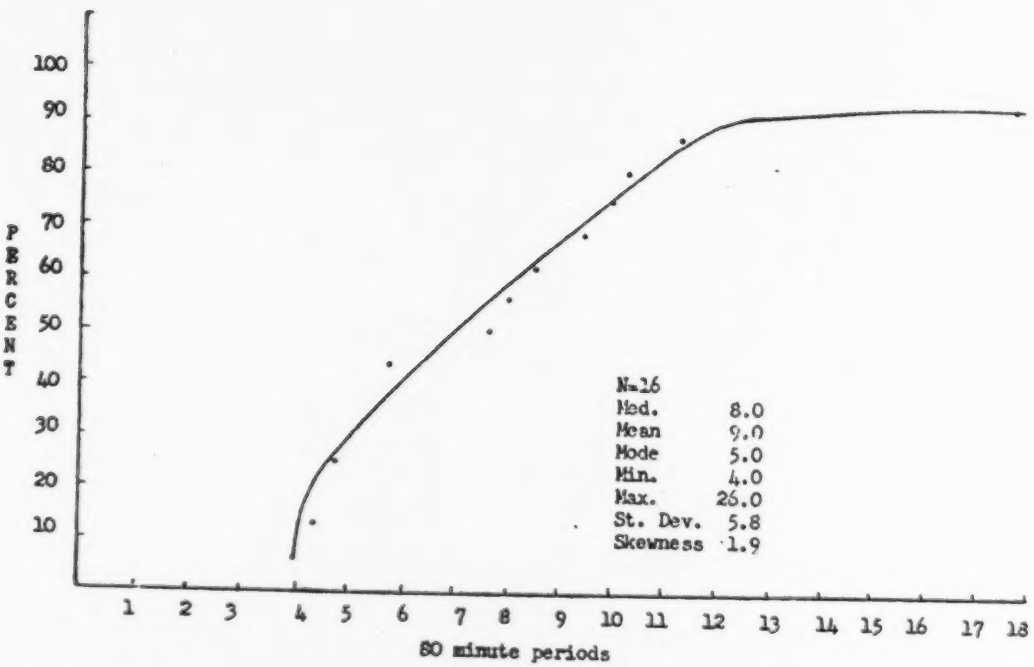


Figure 10.—Cumulative frequency distribution of time to completion, IMU C1.

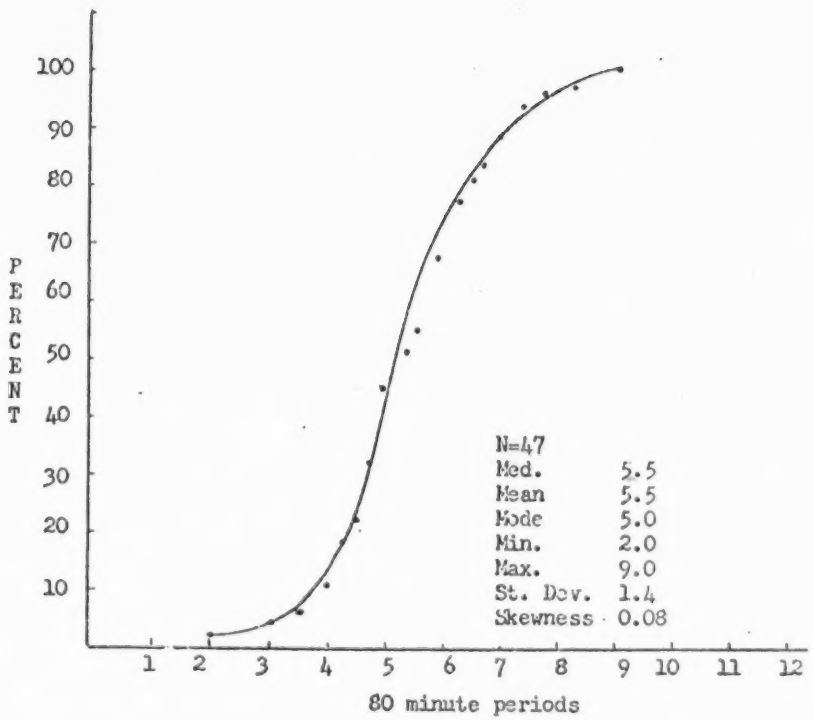


Figure 11.—Cumulative frequency distribution of time to completion, IMU C2.

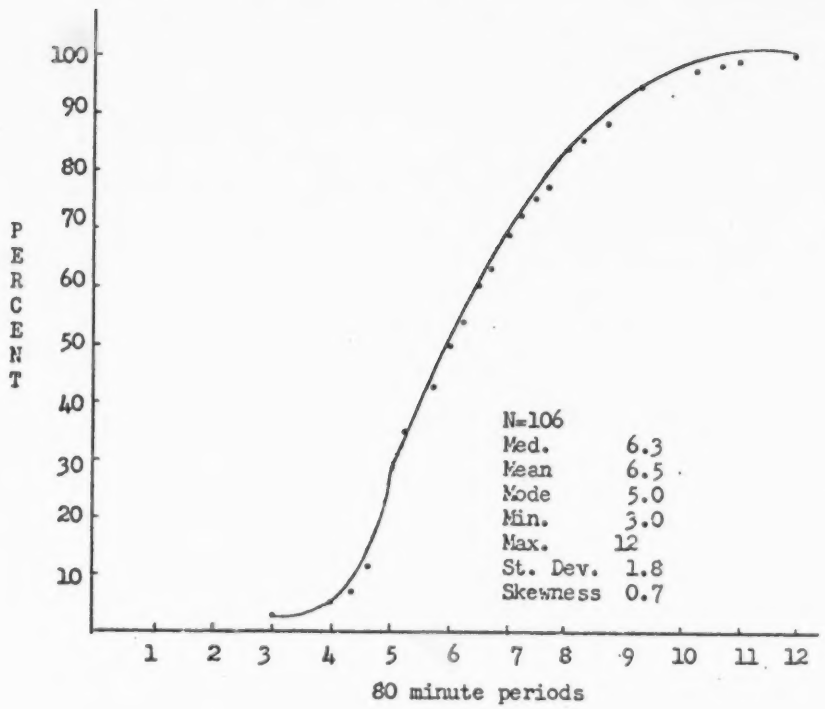


Figure 12.—Cumulative frequency distribution of time to completion, IMU C3.

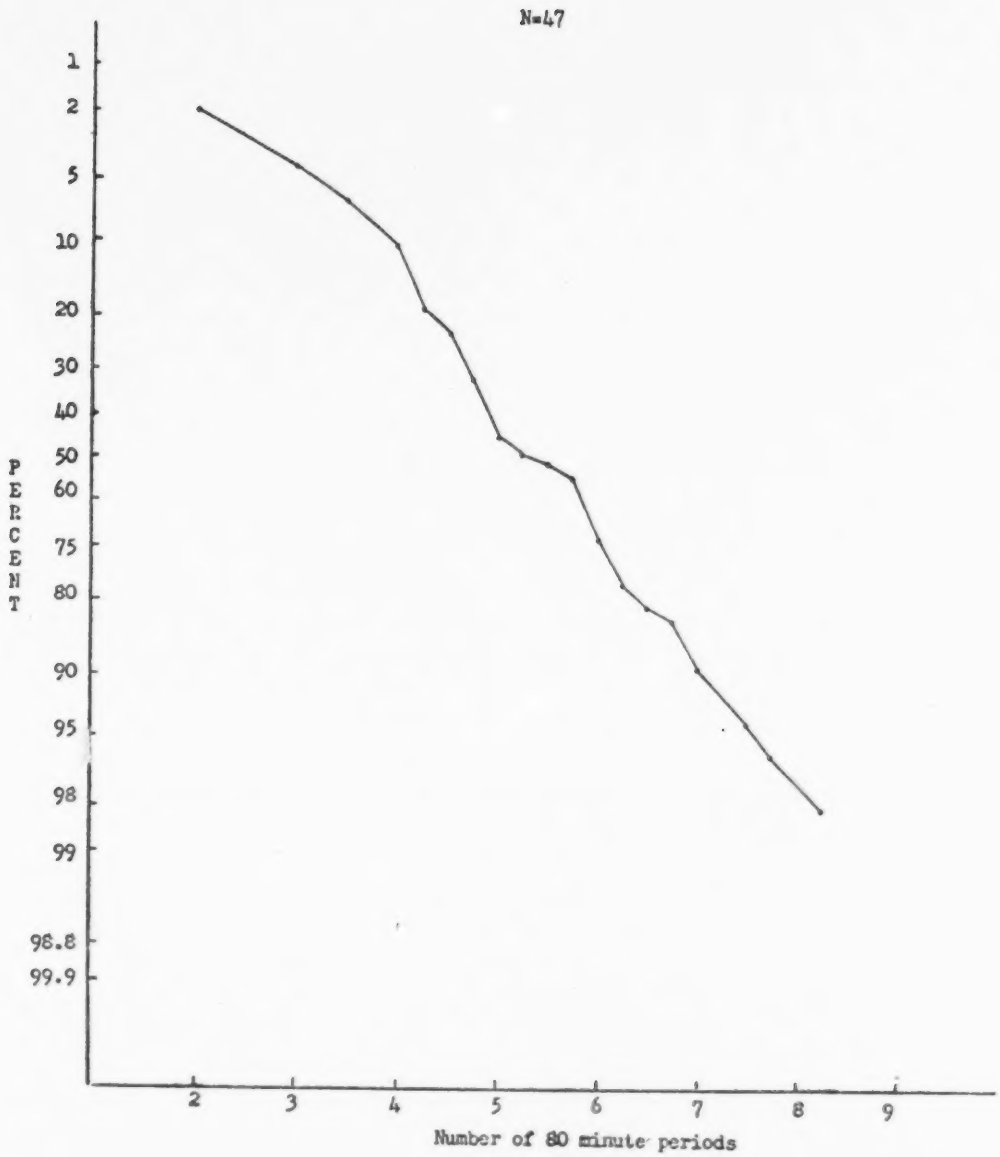


Figure 13.—Cumulative frequency distribution of time to completion, IHU C2, on a normal probability scale.

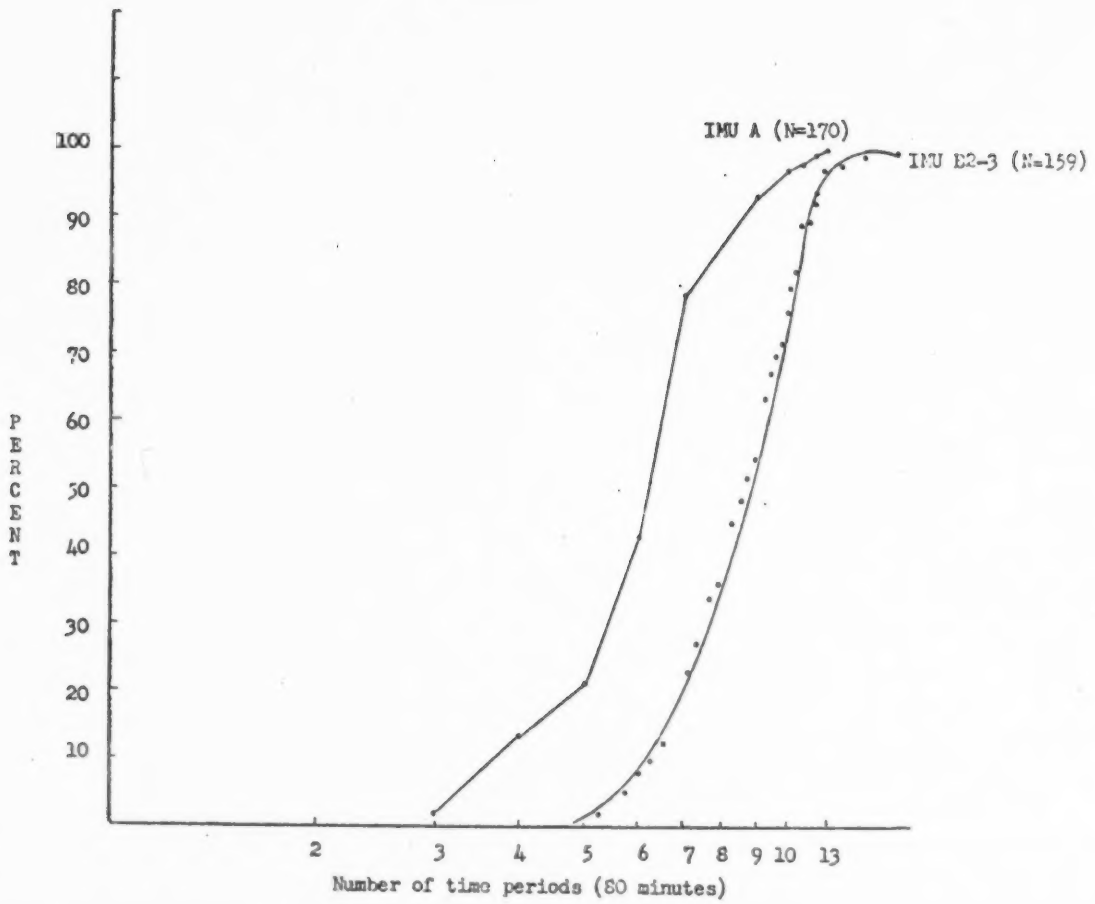


Figure 14.—Cumulative frequency distribution of time to completion, IMU A and IMU B2-3, on a semilog scale.

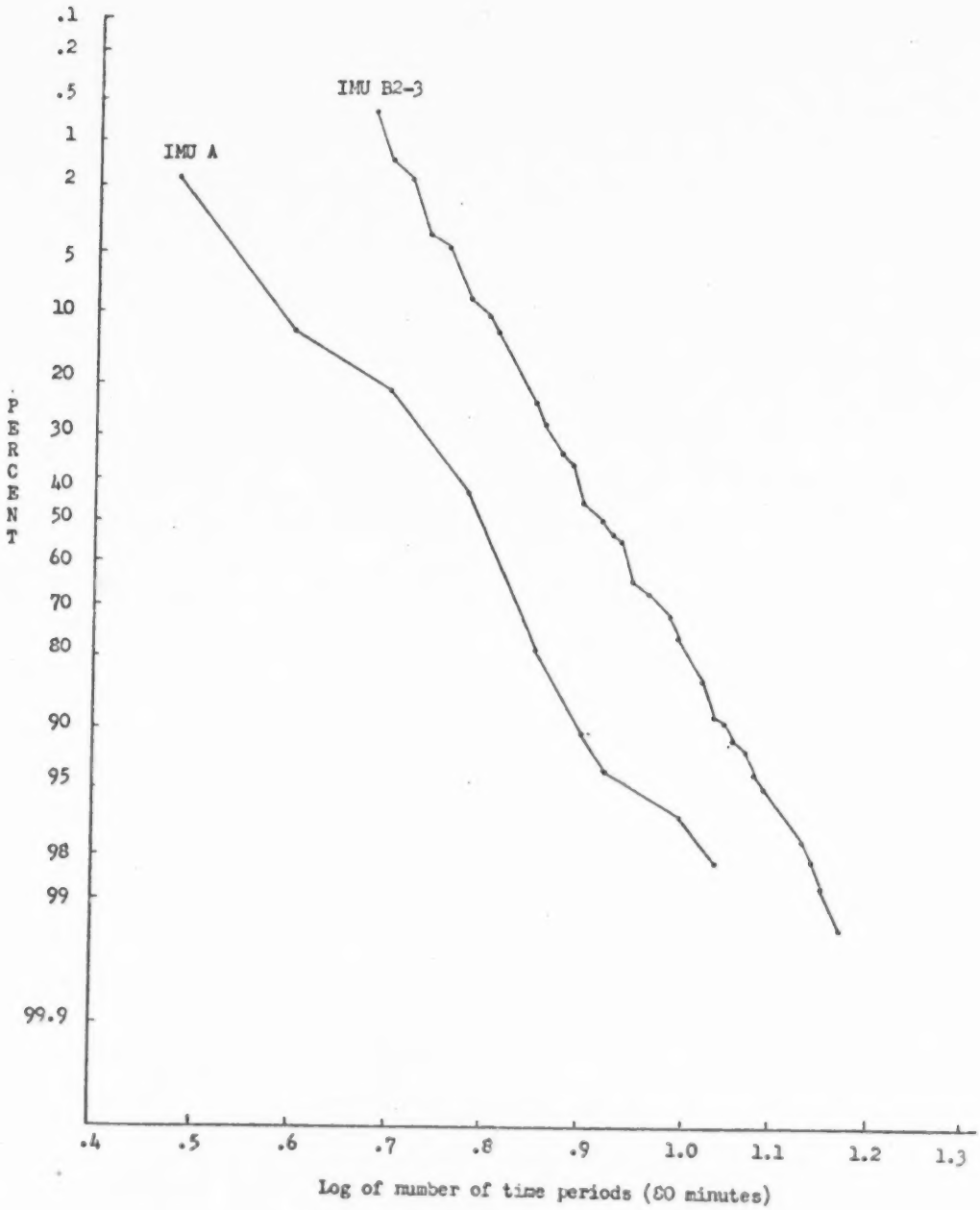


Figure 15.—Cumulative frequency distribution of time to completion, IMU A and IMU B2-3, on a normal probability scale.

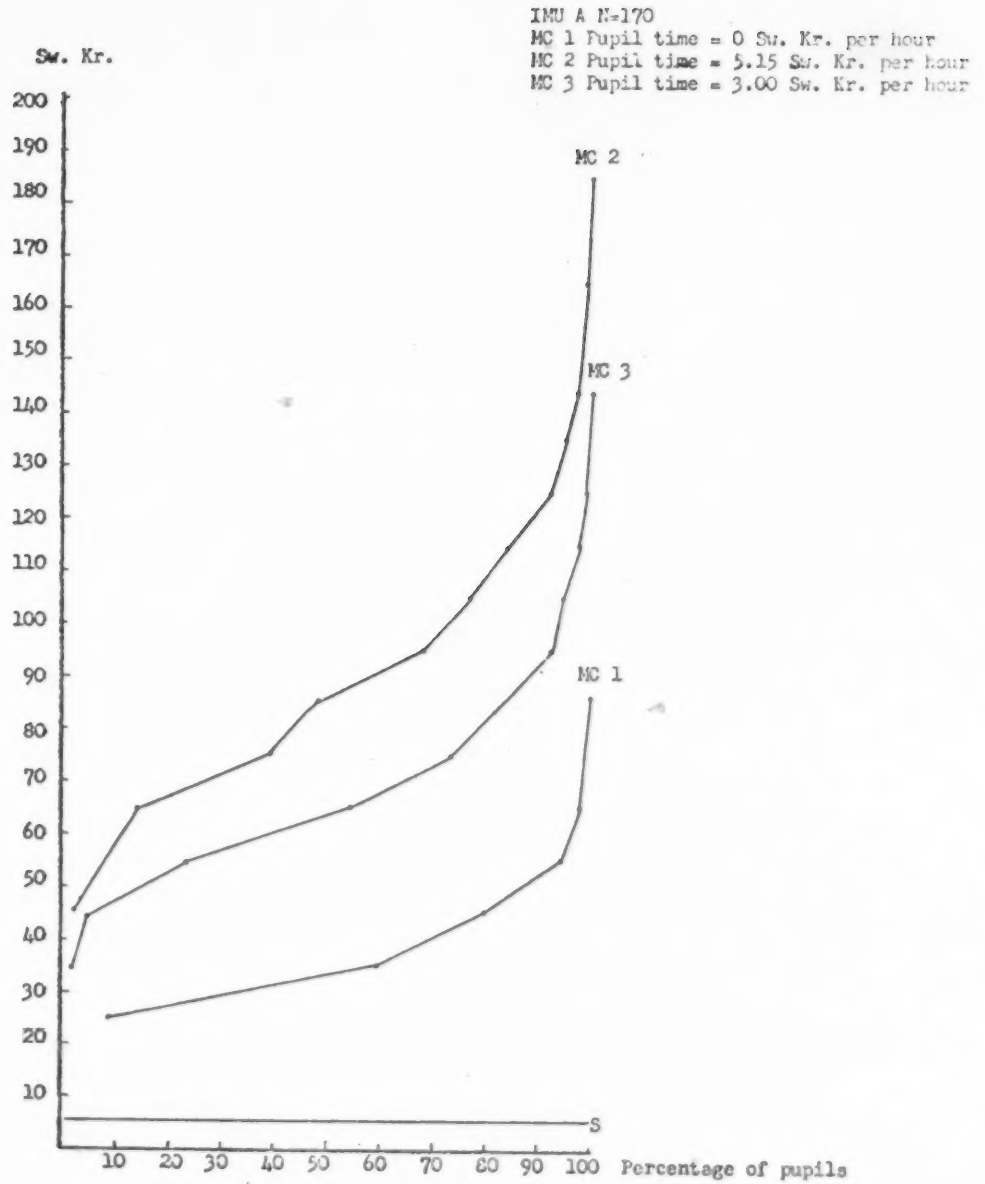


Figure 16.—Marginal cost of step IMU A, S = cost of materials.

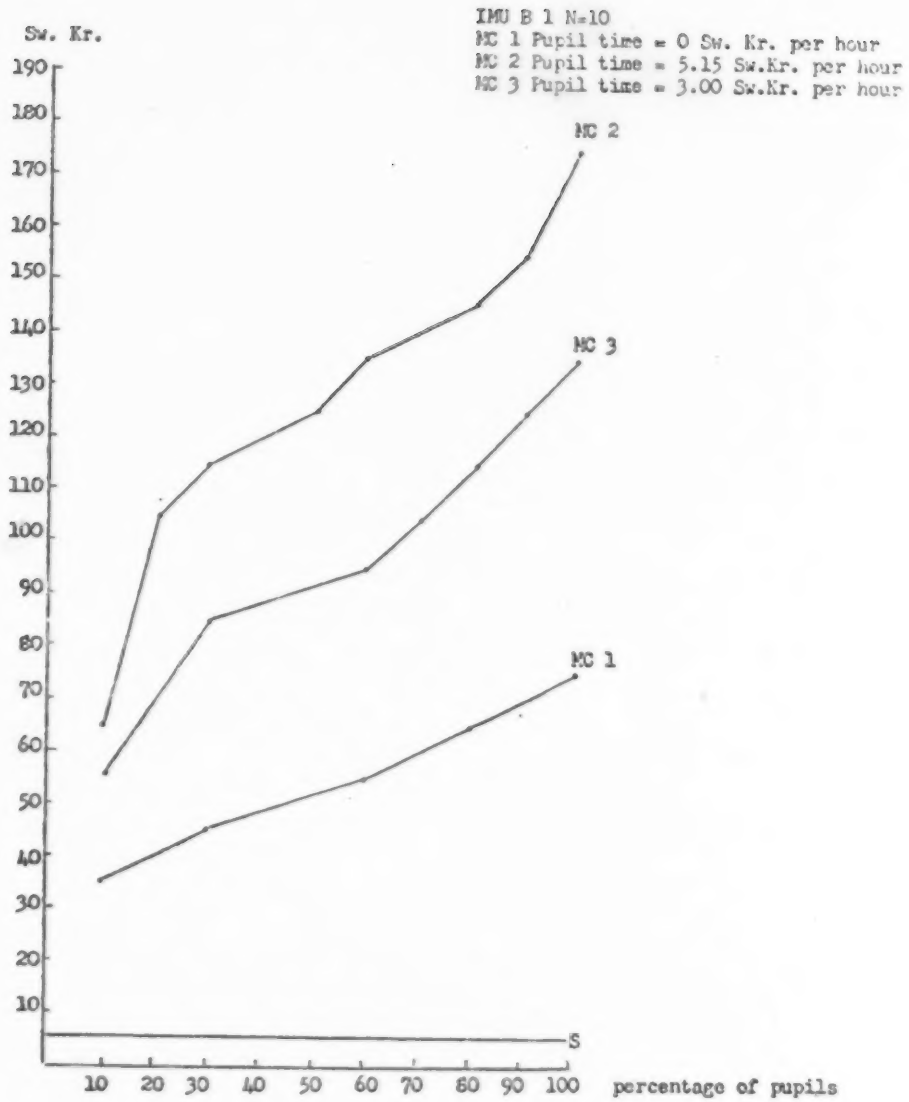


Figure 17.—Marginal cost of step IMU B1, S = cost of materials.

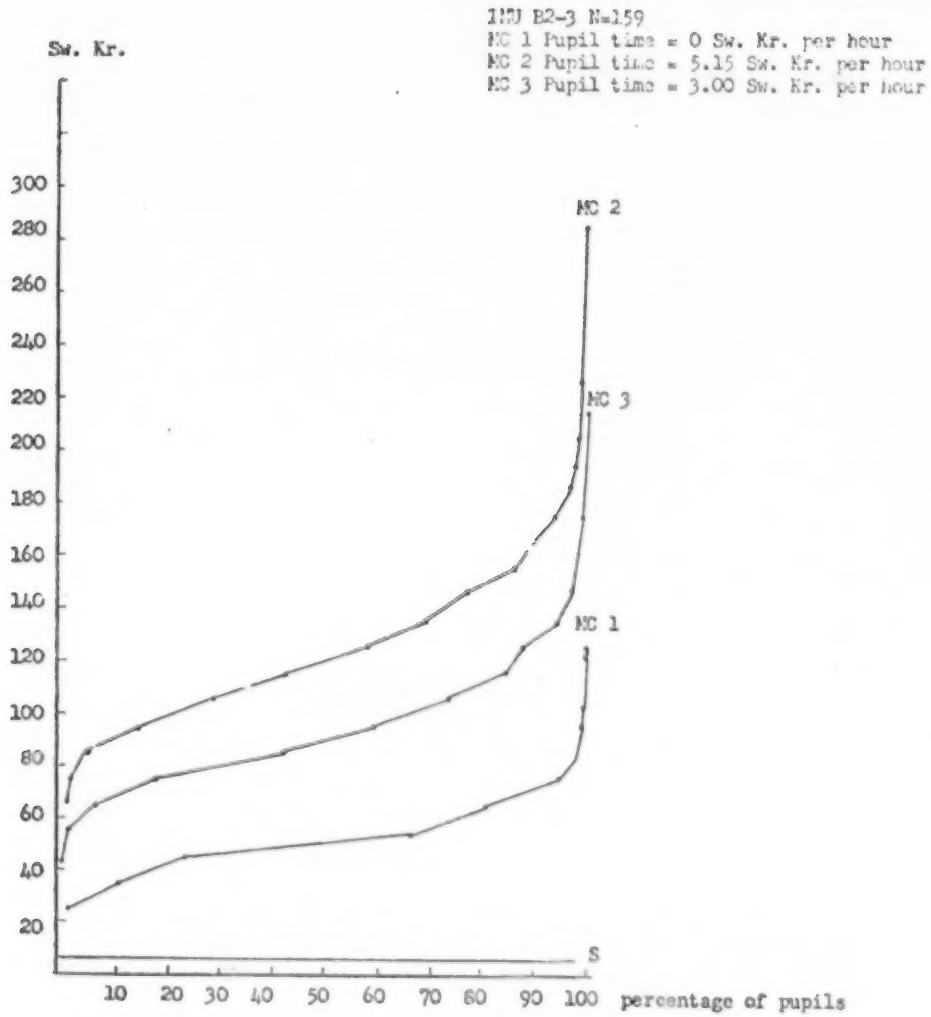


Figure 18.—Marginal cost of step IMU B2-3, S = cost of materials.

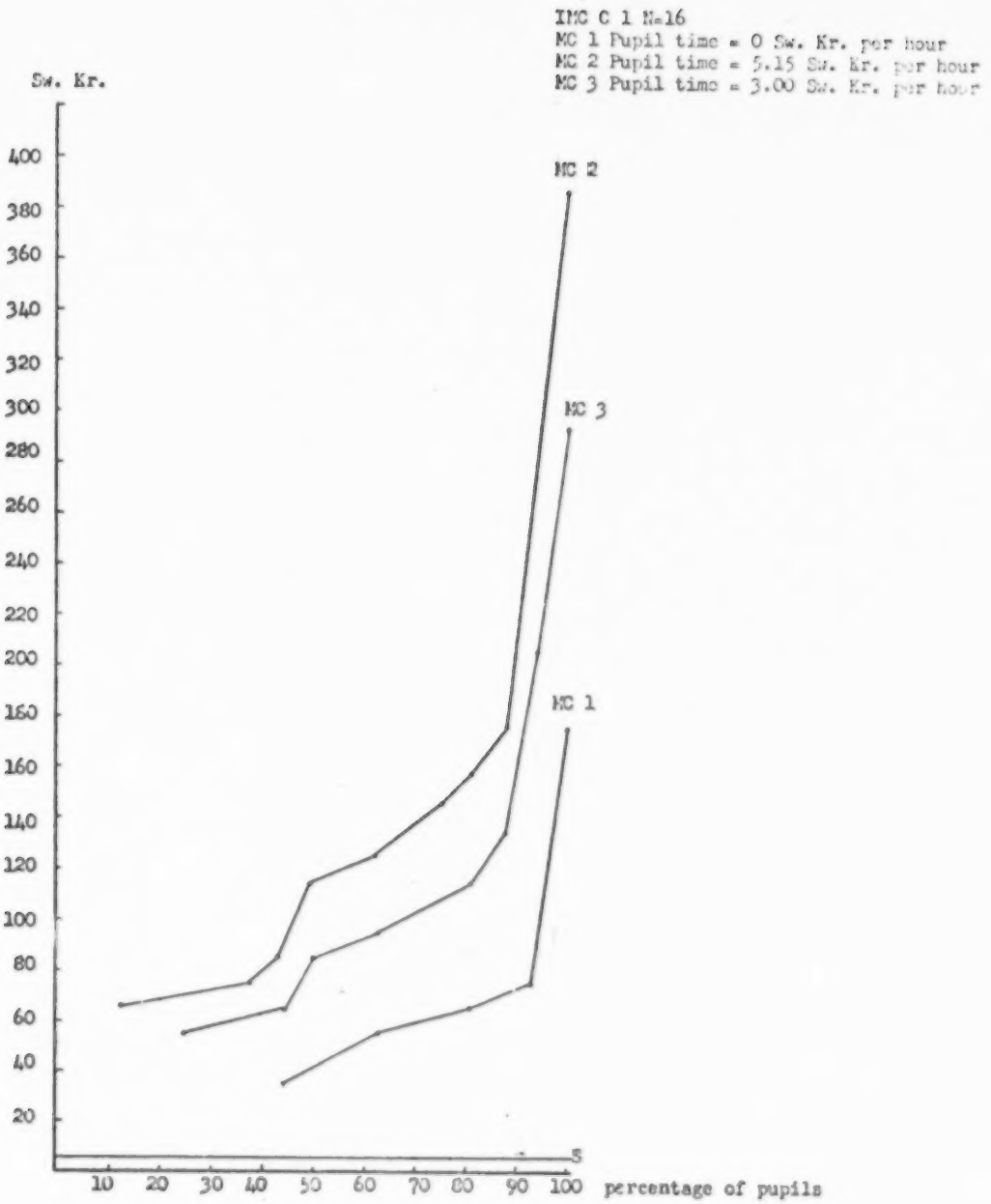


Figure 19.—Marginal cost of step IMU C1, S = cost of materials.

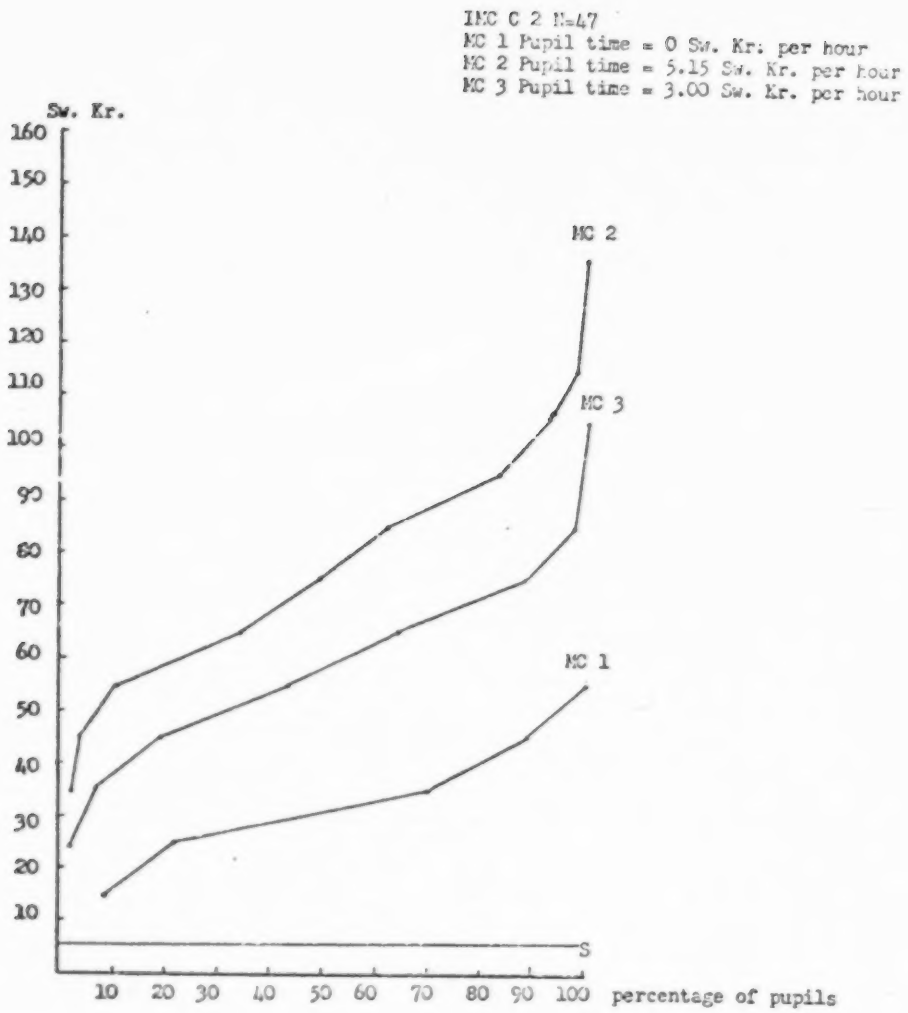


Figure 20.—Marginal cost of step IMU C2,S = cost of materials.

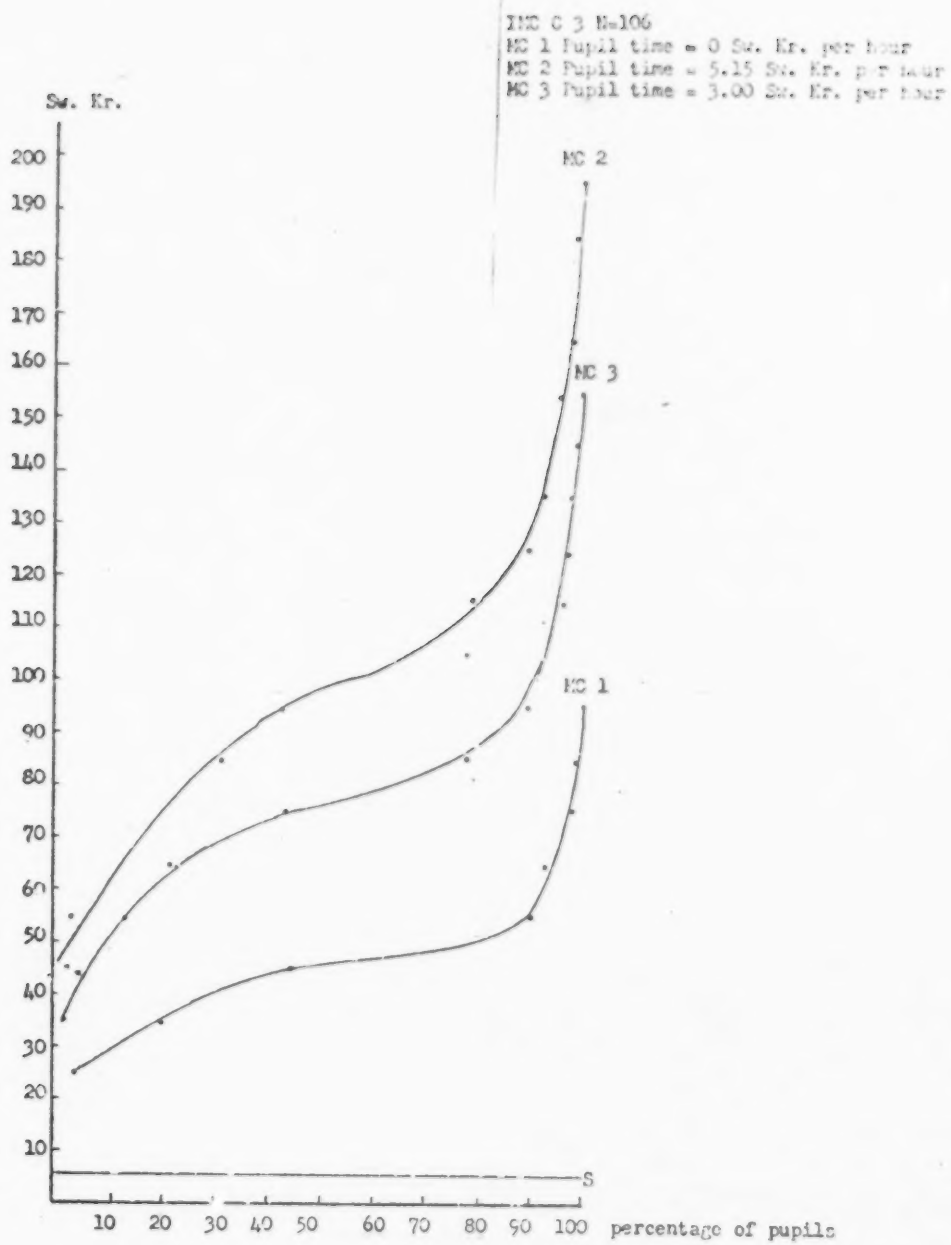


Figure 21.—Marginal cost of step IMU C3, S = cost of materials.

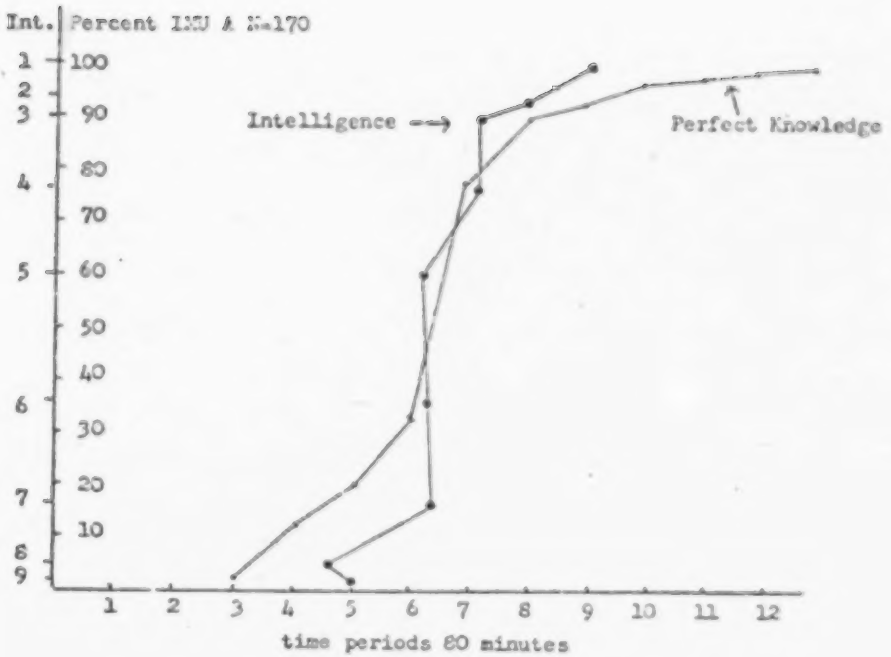


Figure 22.—Cumulative frequency distribution as predicted by "perfect knowledge," and by intelligence.

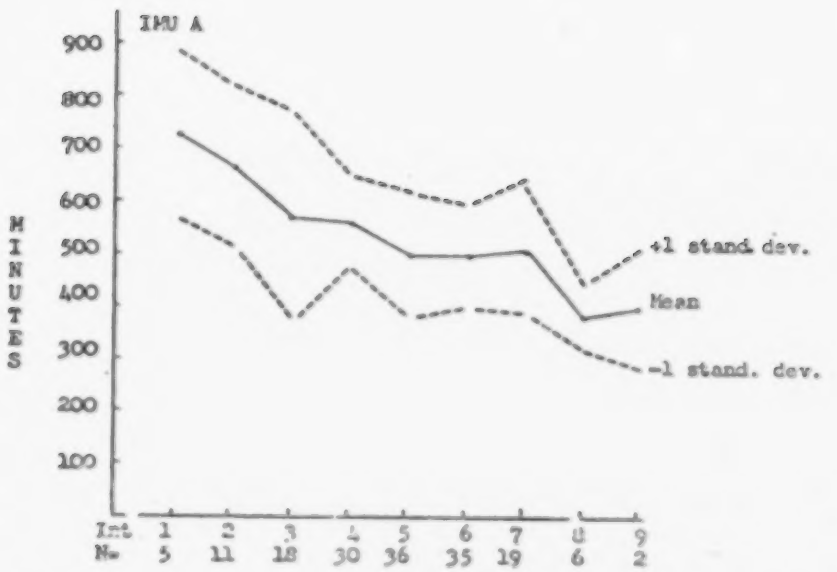


Figure 23.—Mean of time to completion, IMU A, for different intelligence groups (1 low - 9 high)

CHAPTER VIII

CONCLUSION

The major aim of this study was to develop cost-benefit methods relevant to timing in education and applicable to educational decision making. Output to be maximized was defined as skill learning, although it was fully acknowledged that schools are in fact expected to "produce" other kinds of output as well.

After a review of relevant learning theory and a discussion of input and output variables in education, a marginal cost model was developed (Chapter III), tracing costs from differential time that students need in order to reach mastery on a specified learning task. Various aspects of this model were then discussed and related to mastery learning theory. In particular we were trying to find out what some broadly stated societal goals of "equalizing" would mean within the limits of the presented cost model.

A multistep approach was then introduced (Chapter VI); by analyzing a simple sequence of steps, one step building upon the preceding ones, we were able to draw some implications with respect to the costs of successive steps. The amount of time spent in learning at one level, it was assumed, could influence time needed at later levels. The programming model, which was laid out in the same chapter, generalized the idea of time allocation so that we could analyze conditions for optimal allocation of not only student time, but also, and simultaneously,

the optimal use of other resources as well, such as teacher and school space time. The objective function was a learning-production function expressed as number of students taking specified steps, that were either constructed or weighted to be equivalent and hence comparable. Restrictions on production were resources at disposal, expressed in terms of "time budgets" and a monetary budget.

Finally the theoretical construct was put to a partial test. The empirical analysis was based on time-to-completion data taken in a non-experimental learning sequence in mathematics. The students were allowed to take the time they needed to complete a step before going on to the next level. Data on time required for learning a sequence of mathematics tasks, and on mathematics ability, intelligence, marks in school subjects, and family background were obtained for all students in grade 7 in one school (13 year olds mostly; grade 7 is the first year of the upper department of the Swedish comprehensive nine-year elementary school.) The purpose of the empirical study was twofold:

1) to estimate empirical MC curves in order to illustrate the basic cost model developed in Chapter III;
2) to throw some light on the possible determinants of students' time to completion.

The first task is quite straightforward, involving nothing more complicated than the plugging in of economic data concerning teacher cost, costs of books and materials, and so on, even though, as might be expected, there were some difficulties in arriving at accurate figures for some items, this being especially true for the opportunity cost of students' time.

Time to complete the first step (IMU A) was shown to correlate moderately with other measures of school achievement and with intelligence but not with family background (family income and father's occupation). For the second step (IMU B) the correlation with measures of intelligence and achievement was low and for the third step (IMU C) it had disappeared. Two factors may contribute to this decline in correlations; first homework is introduced with the second step, and second, students select among options after the first step and those options differ in difficulty. Some of the most powerful of the determinants of individual variation in time spent in learning evidently were not captured by the data collected, but such determination was not the main purpose of this study. The time distributions are important in themselves, regardless of what places individuals at various positions in those distributions. A major finding of the empirical research was the positive skewness of time distributions, resulting in a sharp upturn of the MC-curves, approximately for the ten per cent slowest students. Whether this is a general trait or typical of this particular sample only, or maybe typical of learning mathematics, one cannot tell.

This and other studies show clearly the dilemma when trying to use the time variable in order to "equalize" learning, whatever meaning we would like to attach to this equalization. If the tentative conclusion of this research is correct, that typically time distributions are positively skewed, this fact would seem to aggravate the problem. The slow learner will be left more and more behind, that is the important fact of the matter; and those students

who are in the skewed part of the distribution will be especially outstanding in this negative sense. Often it is asserted that in learning situations where time individualization is provided for, competition between students would be minimized and the students would presumably "compete with themselves" or with a fixed standard. Is it not possible, though, that the slow learner will feel "out-competed" by being left behind, with the same detrimental effects as are assumed from competition in more traditional learning situations.

A possible consequence, therefore, of introducing mastery learning or other types of instruction allowing students to take differential time, might be that the society would want to minimize time differences, and thus aim certain measures towards the slow students. In an economic analysis of such measures the marginal cost curve of this study should be of some value, since we here have to compare two forces working on the margin and in opposite directions: a downward push since time for the slow students will be lowered (if measures are successful) and an upward pull by increased costs for those students. If present day school practices are technologically inefficient, as is often assumed, this cost increase might be damped, but insofar as equalization is the goal increase in cost could hardly be avoided. The possibility of over-all increases in efficiency (at all levels and for all, or most, students) is another matter, which introduces the much larger question of how to use resources "freed" by increased efficiency.

As pointed out in Chapter I, learning is a lifelong process beginning before the child enters school and extending beyond the end

of formal schooling. Applying our concept of step sequences to this broader context, we may speculate on possible out-of-school determinants of time-to-completion patterns in school subjects, and on any further effects beyond school. By widening the scope of the analysis we do, in effect, examine the efficient use of time over a whole lifecycle, placing school learning in a context of what precedes, accompanies, and comes after schooling. It is evident, then, that one possible explanation of the large variability between individuals in their time to complete a learning task in school may be that for some individuals the learning of pre-school steps is inadequate as an input in later steps, taken in school. This lack of preparation may, furthermore, accumulate with each step taken, if no remedy is provided. In this sense the present analysis ties in with the whole complex of problems related to deprived living, and the costs of minimizing time differentials may be looked upon as compensatory outlays.

Similarly we may extend the analysis forward and speculate about the possible results of learning in school when this learning becomes an input in post-school activities. The adequacy of this input will no doubt have a significant influence not only directly on the efficiency of an individual's use of time at work and his possibilities of enjoying leisure, but also on further learning through experience, whether informal or formalized on-the-job training. So once again we touch upon the problem of deprivation. In this broad context the opportunity cost of not taking remedial measures, if they exist, could be very high indeed.[1]

---

[1]For a thorough discussion of the concept of opportunity costs in

The programming model gives some clues as to how resources may be reallocated in order to be used more effectively. A good teacher may behave as if maximizing output in a way that comes close to what the programming model describes, allocating her time and other inputs under her command so as to equalize marginal products. Other teachers may have to be instructed to do so.

Intuitively it might seem that the able students, who are in most cases also the fast learners are the ones who would benefit mostly from an additional unit of any input. Intuition may, however, be very misleading; whereas the fast and able learners certainly show higher average learning products with respect to inputs, this may not be true for marginal products. Some of the slow learners are, no doubt, limited by poor genetic endowment, others, however, may be hindered by lack of relevant prior learning, at home and/or in earlier schooling. In economic terms: prior investment in the human capital of this category of slow learners has been low relative to other students. In this latter category we might get relatively high marginal returns to additional investment, at least if measures to speed up their learning are taken early enough during the course of their schooling.

The programming model, although operational in the conceptual sense was not empirically tested and is, furthermore, not even testable as it is now formulated. This study seems to suggest that future

---

general and as related to investment in human capital in particular, see: Mary Jean Bowman, "The Costing of Human Resource Development," in The Economics of Education, ed. by E.A.G. Robinson and J.E. Vaizey (New York: St. Martin's Press, 1966), pp. 421-450.

research be directed towards this problem. That would, incidentally, relate very neatly to work, now being done in economics of education, directed towards estimation of school production functions. It is clear that among the empirical problems associated with the programming model is a probable, non-linearity of the objective function. The constraints are linear and hence in a form that is easily applicable empirically. In order for the objective function is to be estimated, we must, in the present "state of the art," be able to put it in a linear or quadratic form.

The theoretical conclusions that were derived from the programming model would, however, seem to justify its existence, even though it might not be fully explored empirically. This, after all, is a characteristic the model shares with other economic theory. The most important contribution of the programming model to an economic theory of education seem to be the identification of marginal product of student time with learning rate and a shadow pricing of this time resource that makes price proportional to learning rate, hence inversely proportional to time spent in learning.

The immediate practical value of the present study seems to be that it offers a tool for economic evaluation of different teaching methods by an analysis of the costs of time in learning. We have already, during the course of exposition, pointed out some considerations for educational decision making when applying the marginal cost model, of which some of the most important can be summarized:

1) We must have a thorough knowledge of the distribution of time-to-completion for the target population of students, who may or may not comprise a whole age cohort.

2) In addition to knowing the time curve we must in some applications of the model also be able to identify individual students, or rather types of students, along this curve.

3) Some students will, inevitably, overperform with respect to a criterion. This must somehow be accounted for in a final analysis.

4) We also assume that we have accurate cost data. Difficulties of measurement and the cost-benefit approach were reviewed in Chapter III.

It is to be expected that when going from this oversimplified model of the world to reality itself, we will have to make some adjustments. We assumed, for example, that all students would eventually reach mastery on a given step. In real life we will have to deal with some failing students, no doubt, and the costs of those must also count in an economic evaluation. They contribute to costs but not to benefits.

The present-value-model of the costs of sequential steps has the double virtue of being directly applicable in a straightforward way and of supplying the decision maker with a large amount of information. Especially valuable is the possibility we have of using the model to control long term effects. These are, of course, the ones we usually strive for in education; yet all too often it turns out that spectacular new teaching methods do not have long lasting effects. When we are using education to counteract cultural deprivation, for

example, such a neglect may be tragic and run counter to the long-ranging goals society has set for itself with respect to welfare.

An ever growing proportion of the population in most countries, technologically developed or not, are engaged in formal education—whether as students, teachers, or other school personnel. The concern of the present study is not with the very complicated problem of allocation in the large of societal resources to education. The question is rather, put in one brief sentence: How do we make the best use of those resources once they are entrusted to the educational sector? It was in the spirit of perhaps being able to throw some light on some of the problems involved here, that this study was undertaken.

SELECTED BIBLIOGRAPHY

Adelman, Irma and Parti, Michael. "The Determinants of Student Achievement: A simultaneous Equation Approach." Unpublished manuscript. Chicago: Northwestern University, 1971.

Anastasi, Anne. Differential Psychology. Individual and Group Differences in Behavior. New York: MacMillan, 1958.

_____. "Heredity, Environment and the Question 'How?.'" Psychological Review, LXXII, (1958), 197-208.

Anderson, C. Arnold and Bowman, Mary Jean. "Theoretical Considerations in Educational Planning." Economics of Education I. Edited by Mark Blaug. Penguin Modern Economic Readings, Harmondsworth, England: Penguin Books, 1969.

Anthony, Bobby C. M. "The Identification and Measurement of Classroom Environmental Process Variables Related to Academic Achievement." Unpublished Ph.D. dissertation, University of Chicago, 1967.

Becker, Gary S. "A Theory of the Allocation of Time." Economic Journal, LXXV, No. 299 (September, 1965), 493-517.

_____. Human Capital. New York: Columbia University Press, 1964.

Behr, Merlyn J. "A Study of Interactions between 'Structure of the Intellect' Factors and Job Methods of Presenting Concepts of Modulus Seven Arithmetic." Unpublished Ph.D. dissertation, Florida State University, 1967.

Block, James H., ed. Mastery Learning Theory and Practice, with selected papers by Peter W. Airasian, Benjamin S. Bloom, John B. Carroll. New York: Holt, Rinehart and Winston, Inc., 1971.

_____. "The Effects of Various Levels of Performance on Selected Cognitive, Affective, and Time Variables." Unpublished Ph.D. dissertation. University of Chicago, 1970.

Bloom, Benjamin S. "Learning for Mastery," Evaluation Comment, Vol. I, No. 2. Los Angeles: Center for Study of Evaluation of Instructional Programs, University of California, 1968.

_____. Stability and Change in Human Characteristics. New York: John Wiley and Sons, Inc., 1964.

Bloom, Benjamin S., Davies, Allison and Hess, Robert. Compensatory Education for Cultural Deprivation. New York: Holt, Rinehart and Winston, Inc., 1968.

Bowles, Samuel. Planning Educational Systems for Economic Growth. Cambridge, Mass.: Harvard University Press, 1969.

_____. "Towards an Educational Production Function." Education, Income, and Human Capital. Edited by W. Lee Hansen. Studies Income and Wealth. New York: Columbia University Press, 1970.

Bowles, Samuel S. and Levin, Henry M. "The Determinants of Scholastic Achievement: An Appraisal of Some Recent Findings." Journal of Human Resources, III, No. 1 (Winter, 1968), 3-24.

Bowman, Mary Jean. "Converging Concerns of Educators and Economists." Comparative Education Review. VI, No. 2 (October, 1962), 111-119.

_____. "Schultz, Denison and the Contribution of 'Eds.' to National Income Growth." Journal of Political Economy, LXXII, No. 5 (October, 1964), 450-464.

_____. "Social Returns to Education." International Social Science Journal, XIV, No. 4 (1962), 647-659.

_____. "The Costing of Human Resource Development." The Economics of Education. Edited by A. E. G. Robinson and J. E. Vaizey. New York: St. Martin's Press, 1966.

Bowman, Mary Jean and Anderson, C. Arnold. "Distributional Effects of Educational Programs." Income Distribution Analysis. Report from Conference on Income Distribution Analysis with Special Reference to Problems of Rural People. Raleigh, N.C.: North Carolina State University, June, 1966.

_____. "Relationships among Schooling, Ability and Income in Industrialized Societies." Bildungsökonomie—Eine Zwischenbilanz. Economics of Education in Transition. Edited by Klaus Hufner and Jens Nauman. Stuttgart: Ernst Klett Verlag, 1970.

Burkhead, Jesse. Input and Output in Large City High Schools. Syracuse, New York: Syracuse University Press, 1967.

Carroll, John B. "A Model of School Learning." Teachers College Record, LXIV (May, 1963), 723-733.

Coleman, James S. et al. Equality of Educational Opportunity. Washington: U.S. Office of Education, 1966.

Cronbach, Lee J. "How Can Instruction be Adapted to Individual Differences," in Learning and Individual Differences. Edited by Robert M. Gagne. Columbus, Ohio: C. E. Merrill, 1967.

Cronbach, Lee J. and Snow, R. E. Individual Differences in Learning Ability as a Function of Instructional Variables. Final Report, USOE, Contract No. OEC 4-6-061269-1217. Stanford, California: Stanford University, 1969.

Dahllöf, Urban S. Ability Grouping, Content Validity and Curriculum Process Analysis. New York: Teachers College Press, 1971.

_____. Skoldifferentiering och undervisningsförlopp. Komparativa mål och process-analyser av skolsystem I. Göteborg Studies in Educational Sciences 2. Stockholm: Almqvist och Wiksell, 1967.

Denison, Edward F. The Source of Economic Growth in the United States and the Alternatives Before Us. New York: Committee for Economic Development, 1962.

_____. Why Growth Rates Differ. Postwar Experience in Nine Western Countries. Washington, D.C.: Brookings Institution, 1967.

Fleishman, Edwin A. "Individual Differences and Motor Learning." Learning and Individual Differences. Edited by Robert M. Gagné. Columbus, Ohio: C. E. Merrill, 1967.

Gagné, Robert M., and Paradise, Noel E. "Abilities and Learning Sets in Knowledge Acquisition." Psychological Monographs, Vol. LXXV, No. 14 (Whole No. 518, 1961).

Gagne, Robert M. Basic Studies of Learning Hierarchies in School Subjects Final Report, Project No. 6-2949, Contract No. OEC-4-062949-3066. Berkeley, California: University of California Berkeley, 1970.

_____. The Conditions of Learning. New York: Holt, Rinehart and Winston, Inc., 1967.

Hadley, G. Nonlinear and Dynamic Programming. Reading, Mass.: Addison-Wesley Publishing Co., Inc., 1964.

Hanushek, Eric Alan. "The Education of Negroes and Whites." Unpublished Ph.D. dissertation, Harvard University, 1968.

Husén, Torsten. Talent Opportunity and Career. Stockholm: Almqvist och Wiksell, 1969.

Katzman, Martin T. "Distribution of Production in a Big City Elementary School System." Yale Economic Essays. Vol. VIII, No. 1 (Spring, 1968).

Kiesling, Herbert J. "Measuring a Local Government Service: A Study of School Districts in New York State." The Review of Economics and Statistics, XLIX, No. 3 (August, 1967), 350-358.

Kim, Hogwon. "Learning Rates, Aptitudes and Achievements." Unpublished Ph.D. dissertation, University of Chicago, 1968.

Lancaster, Kelvin. Mathematical Economics. New York: The MacMillan Company, 1968.

Levin, Henry M. "A New Model of School Effectiveness." Do Teachers Make A Difference? A Report on Recent Research on Pupil Achievement. U.S. Department of Health, Education and Welfare. Washington, D.C.: U.S. Government Printing Office, 1970.

Marklund, Sixten. Skolklassens Storlek och struktur. Stockholm: Almqvist och Wiksell, 1962.

Mincer, Jacob. "On-the-Job Training: Costs, Returns, and Some Implications." The Journal of Political Economy, LXX, No. 5, part 2 (Supplement, October, 1962) 9-79.

Prest, A. R. and Turvey R. "Cost-Benefit Analysis: A Survey." Surveys of Economic Theory, Volume III, Resource Allocation. American Economic Association, Royal Economic Society. London: MacMillan and Company Limited, 1966.

Schultz, Theodore W. The Economic Value of Education. New York: Columbia University Press, 1963.

Swedner, Harald and Edstrand, Gisela. "Skolsegregation i Malmö." Mimeographed. Lund: Department of Sociology, University of Lund, 1969.

Thomas, J. Alan. "Efficiency in Education: A Study of the Relationship Between Selected Inputs and Mean Test Scores in a Sample of Senior High Schools." Unpublished Ph.D. dissertation, Stanford University, 1962.

_____. The Productive School. New York: John Wiley and Sons, Inc., 1971.

Windham, Douglas M. Education, Equality and Income Redistribution. Lexington, Ma.: D. C. Heath and Company, 1970.

Department of
Educational and
Psychological Research
School of Education
Malmö, Sweden

194

Abstract card

Christoffersson, N.-O. The economics of time in learning. Didakometry (Malmö, Sweden: School of Education), No. 34, 1971.

In the economics of education a main concern is with the study of relationships between educational objectives or goals and the scarce resources used as means towards those ends. This study is focused on the cost side; it treats student time in learning as one of those scarce resources and relates costs to time in education. The empirical data were taken in a learning situation, using the IMU system a method of teaching mathematics, developed at the Malmö School of Education. The study was undertaken and completed as a Ph.D. dissertation at the Comparative Education Center, University of Chicago.

Indexed:
1. Comparative education
2. Economics of education
3. Individualized instruction

Reference card

Christoffersson, N.-O. The economics of time in learning. Didakometry (Malmö, Sweden: School of Education), No. 34, 1971.